SEMIPARAMETRIC GEOMETRIC METHODS FOR EXTRACTING AND MODELING WHITE MATTER VOLUMETRIC STRUCTURES OF THE BRAIN

by

Gary D. Pack

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

Doctor of Philosophy

(Computer Sciences)

at the

UNIVERSITY OF WISCONSIN–MADISON

2012

Date of final oral examination: 12/11/12

The dissertation is approved by the following members of the Final Oral Committee:
Charles R Dyer, Prof., Computer Sciences
Andrew Alexander, Prof., Medical Physics and Psychiatry
M. Elizabeth Meyerand, Prof., Medical Physics and Biomedical Engineering
Moo Chung, Associate Prof., Biostatistics and Medical Informatics
Vikas Singh, Assistant Prof., Biostatistics and Medical Informatics

To Ann for her encouragement, support, and patience

Acknowledgments

I would like to thank and acknowledge my advisor Chuck Dyer. He has been a valuable mentor and has been instrumental in helping me to expand into the wider arena of cross-disciplinary research. I would also like to thank Andy Alexander for his role in funding me and providing me with the opportunity to collaborate in brain imaging. I would like to thank the balance of my committee, Beth Meyerand, Moo Chung, and Vikas Sing, for their help and comments during the development of this work. Additionally, I want to acknowledge the National Library of Medicine training grant 5T15LM007359.

Finally, I would like to thank the Computation and Informatics in Biology and Medicine Training Program and Louise Pape.

contents

list of figures

Abstract

There are important questions that require an understanding of white matter connectivity in the brain, such as how the brain integrates sensory inputs, how white matter tissue and connections are altered by disease, or how cancer tumors should be excised without destroying healthy tissues. These questions require a global understanding of the geometric nature of white matter tissue as well as the organization of white matter structure in local regions of the brain. This thesis presents novel approaches and methods that simultaneously combine local and global geometry of white matter into a single model.

The methods presented in this thesis are called Semiparametric Geometric Modeling (SGM). The SGM fits a nonlinear manifold to Diffusion Weighted Magnetic Resonance Imaging data and produces a nonlinear coordinate system. Specifically, an SGM simultaneously extracts white matter structures and produces a set of functions that together define a model of the white matter. An SGM produces manifold models of the physical white matter structures. This allows the physical structures to be mapped by a multi-dimensional, nonlinear coordinate system that allows points, curves, surfaces, and volumes to be defined by the manifold model. Associated SGM functions can interpolate to the level of a single neural fiber, reveal the path of nerve fiber bundles, and be used to study the interaction e.g. crossing, touching, bifurcating, of fiber bundles throughout the brain. SGM functions can be used to query the manifold structure, allowing data to be organized so as to enable methods such as Functional Data Analysis to be used for statistical analysis of the data.

Software to build SGMs was implemented and a series of experiments were carried out on Diffusion Weighted Magnetic Resonance Imaging data. The data consisted of control subjects and subjects with autism. An SGM was used to simultaneously extract and model two structures for each subject: a portion of the genu of the corpus callosum and the right corticospinal tract. The SGMs were used to map data from imaging space to curves on the manifold. These data curves were the input for group differential analysis using Functional Data Analysis. Group differences were found, based on these structures that are consistent with results from other sources. However, the results also indicate that the group differences were the result of differences in rates in change in data distributions along the structure rather than simply point-wise differences in data at specific locations.

Chapter 1

Introduction

The brain is the most important and complex organ in the human body. 98% of of neural tissue, specialized for sending command signals to all other organs, is concentrated in the brain. The brain contains over 100 billion neurons with several hundred trillion connections and an estimated storage capacity of 2.5 petabytes, enough capacity to record high definition video 24 hours a day for 300 years. The brain controls or supervises everything from autonomic processes like breathing to walking to deep cognitive functions such as seeing and understanding the world around us.

Conditions or diseases that alter the structure of neural tissue, and the connective network, it creates have a significant impact on human health. The Center for Disease Control estimates 1 in 88 children are affected by autism. The World Health Organization estimates that Schizophrenia affects about 24 million people worldwide. Alzheimer's disease is expected to affect 1 in 80 people worldwide. All of these conditions have been tied to changes in white matter tissue and connectivity of the brain. Cancer also modifies tissue and connectivity. Tumors of the white matter distort, infiltrate or destroy white matter fibers. Treatment requires delicately excising the abnormal tissue while simultaneously minimizing functional deficit and maximizing structural integrity.

Understanding brain structure and neural connectivity is fundamental to

understanding "normal" dynamic cognitive processes as well as for treating pathological conditions. First, there are fundamental questions regarding how the brain functions dynamically. For example, how do parts of the brain work together to create memory, perceive and integrate sensory inputs, process emotions, make decisions, and many, many more? There are many tools, such as Functional Magnetic Resonance Imaging and Electroencephalography, that look at dynamic processes in the gray matter of the brain. Additionally, there are methods such as Transcranial Magnetic Stimulation that can stimulate the brain directly to see how brain function changes.

A great deal about how the brain functions has been discovered without taking into account in any detail how the regions of the brain are structurally connected. The next major step is a good understanding of how these areas are structurally connected and how changes in the organization of the connections as well as changes in the quality of the connections modulate cognitive processes. Fortunately, a relatively new imaging method called Diffusion Weighted Magnetic Resonance Imaging provides the basic information needed for studying large scale organization of white matter *in vivo*.

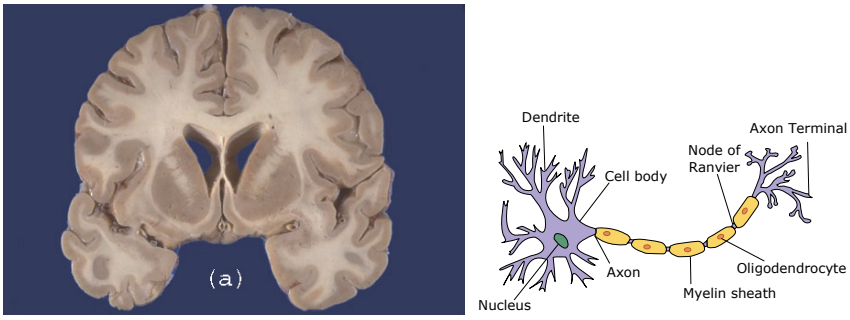

Figure 1.1: (left) Coronal view slice through the mid brain. (right) Sketch of a neuron from [173]. The dark areas are comprised of the cell bodies that make up the gray matter and the light areas are the myelinated axons that make up the white matter.

Diffusion Weighted Magnetic Resonance Imaging (DWI) samples the water diffusivity properties that are determined by the organization of neural tissue. At the

lowest level, the tissue of interest is made up of myelin producing oligodendrocytes and neurons. For simplicity only neural tissue that is involved in long-range communication will be described. The right side of Figure 1.4 shows a sketch of a typical neuron. The cell body and axon terminal are anchored in gray matter tissue. The axon that connects the cell body and axon terminal is wrapped in myelin and propagates signals between them. A single myelinated axon is referred to as a white matter fiber. White matter fibers come in a few different diameters but most are on the order of a few microns in diameter. However, the best resolution currently available in DWI is on the order of cubic millimeters. So, what DWI measures are the tissue level properties of bundles of white matter fibers.

These white matter tissue properties are measured at voxels throughout the volume of the brain and the measurements are organized so that we know which voxels are next to each other. Some of the questions that can't be answered directly from the data in this form are:

1. Where are given fiber bundles are located?
2. How are given fiber bundles organized internally?
3. How do fiber bundles interact structurally with one another?
4. Where do given fiber bundles terminate in the gray matter?
5. Is there a common global organization to the fiber bundles?

Many methods have been proposed to take these voxel level data and construct a map that attempts to answer some of these questions. But until now, no method has integrated all the information available in the data and ideas about how the brain is organized *globally*, from the level of the whole brain (organ) to the regional level to the tissue level. To give some intuition as to the organ level organization, see Figures 1.2 and 1.3. [186] have proposed that white matter fiber bundles are arranged so as to follow the layout of a three dimensional grid that is locally regular but highly distorted at the whole brain level. Since this grid

can then be said to define the geometry of the organization of the fiber bundles. More specifically, the distorted grid is a nonlinear coordinate system on what is known as a nonlinear manifold. The nonlinear coordinate system describes how to “navigate” around the nonlinear manifold. Like any good coordinate system, geometric entities such as points, curves, surfaces, and volumes can be defined in terms of coordinates on the manifold. Manifold properties allow all of the tools and properties of differential geometry, algebraic topology and foliation theory to be brought to bare to analyze and describe the structural organization of white matter in the brain. This thesis presents new methods for building a model of a nonlinear manifold mapped by a nonlinear coordinate system where bundles of white matter connecting different parts of the brain follow the coordinate "axes."

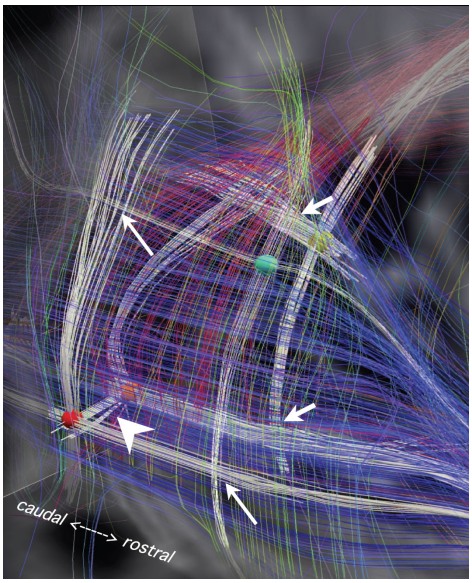

Figure 1.2: Nonlinear geometry of the sagittal stratum in the rhesus occipital lobe. From [186]. Reprinted with permission from AAAS.

A new method is presented in this thesis called Semiparametric Geometric Modeling (SGM). An SGM fits a nonlinear manifold to DWI data, allowing points, curves, surfaces, and volumes to be defined *on that manifold*. Specifically, an SGM produces a set of functions that together define a multidimensional nonlinear

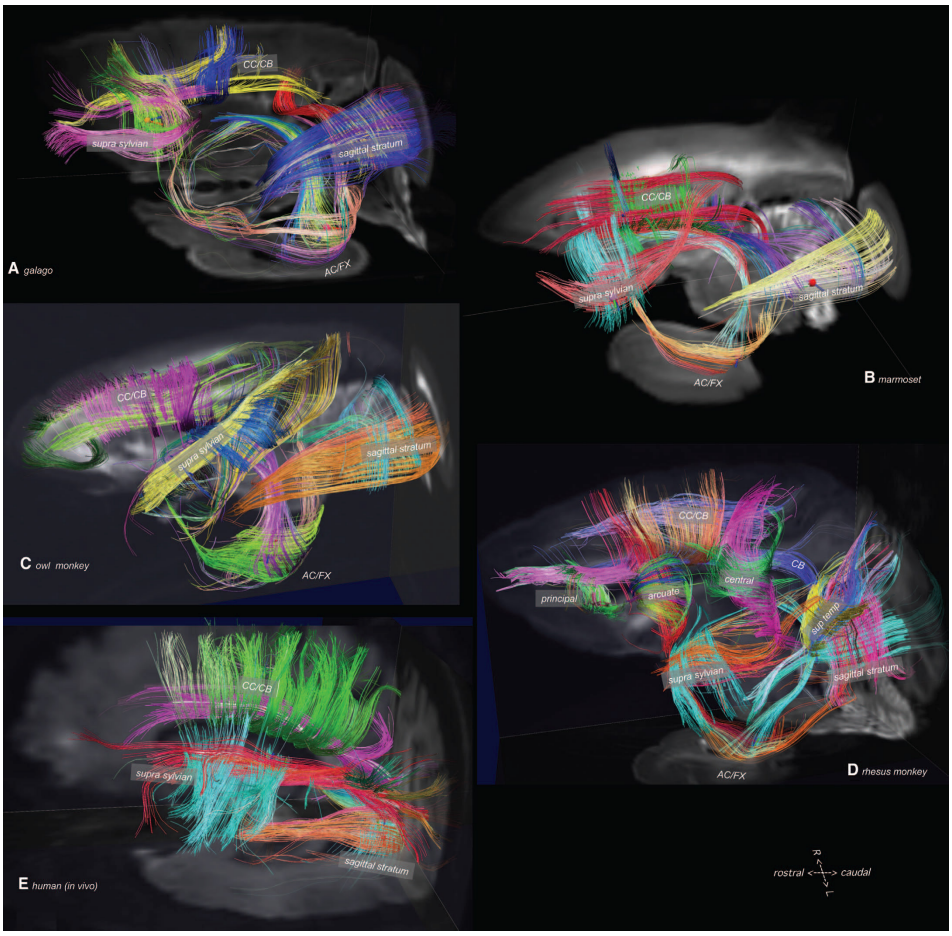

Figure 1.3: Comparative neuroanatomy of white matter organization between (A) galago, (B) marmoset, (C) owl monkey, (D) rhesus monkey, and (E) human. From [186]. Reprinted with permission from AAAS.

coordinate system that maps the nonlinear manifold. The SGM functions can interpolate to the level of a single neural fiber, reveal the path of nerve fiber bundles, and be used to study the interaction (crossing, touching, bifurcating) of multiple fiber bundles throughout the brain.

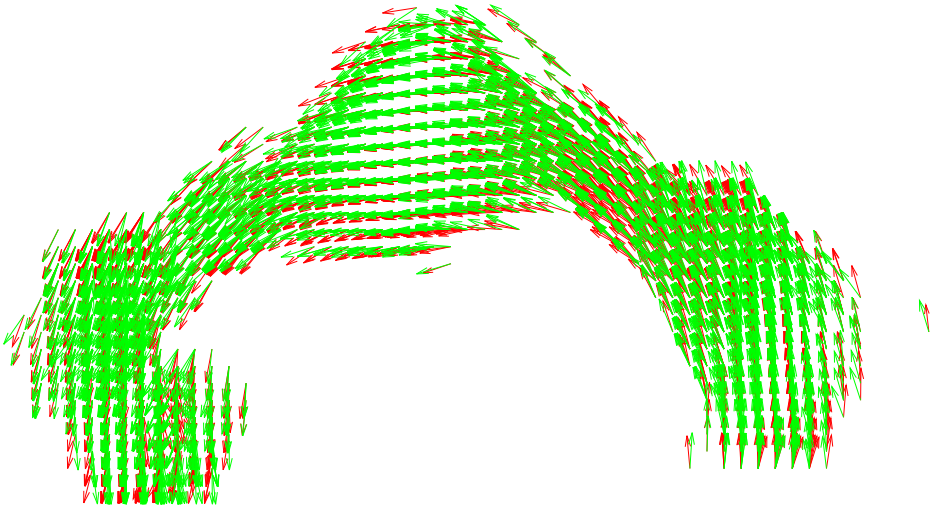

Figure 1.4: The data in this figure are vectors that are tangent to white matter fiber bundles in the genu of the corpus callosum. The green vectors are from the original DWI data. The red vectors were generated from the SGM. The two vector fields have been slightly offset to make visual comparison easier.

1.1 Semiparametric Models

Semiparametric models account for the data at multiple scales simultaneously. They combine global parametric models that can be used to infer large-scale features of a data set but which may occasionally be too smooth to account for small scale features; nonparametric models model data well on a local level but have little control over global structure. Global parametric models have two important properties. First, by accounting for large-scale structure they reduce the complexity and therefore the dimensionality of the data. Second, the form of the parametric model allows us to include prior knowledge about large-scale features in the model in a straightforward manner. Nonparametric models handle local complexity well and allow imposition of prior knowledge about local features of the data. The global parametric models in the SGM are implicit algebraic polynomial functions introduced in the next section and the nonparametric models

are compact kernel functions that describe local geometry alluded to in Section 1.4.

1.2 Implicit Algebraic Polynomial Functions

Implicit algebraic surfaces of the form $S(\vec{x}) = 0$ have long been used in computer vision and computer graphics to describe surface properties of objects. What the equation means is that the surface is determined by all of the points in the domain of $S(\vec{x})$ that evaluate to 0. Two important properties of implicit functions that are central to this thesis are (1) the gradients of implicit functions are normal to the surface defined by the function, and (2) by choosing a range of values other than 0, entire families of implicit functions can be defined. A novel aspect of the SGM is that it fits the gradients of a family of implicit surfaces to a vector field, and that family of implicit surfaces is defined by a single algebraic polynomial.

1.3 Depth Function

When fitting data to a geometric model it is important to decide "where" the model should go. For standard least-squares the criteria for where the model should placed is the most likely part of the data. Given that, the criterion for deciding where the model should be placed is the minimum of the squared residuals. If the model is a line for instance, the squared residual controls the placement and orientation of the line through the data.

In order to fit our SGM to model to DWI data, a similar notion is required. Because of the nature of the problem the squared residual is not appropriate. We turn instead to the notion of statistical depth. When fitting a function it is important for stability and robustness to localize the function near the "center" or deepest point in the data. For point-like data in Euclidean space the geometric median is a robust choice for defining the deepest point in the data. However, for an SGM, the entity that needs to be placed in the deepest part of the data is a surface and

the space it needs to be "centered" in is nonlinear. This dissertation presents a novel depth function called a Geometric Median Surface (GMS) for this purpose. The GMS plays the same role in this nonlinear space that the geometric median does in linear spaces.

1.4 Semi-Supervised Models

The last element of the model is a means of "discovering" the local structure of the data and inferring how it should be added to the global algebraic model. The method chosen for this task is semi-supervised learning. Semi-supervised learning is a classification technique used in machine learning. The basic idea is to take a comparatively small sample of data with known classification "labels" combined with a large amount of unlabeled data and build a classifier that labels all of the data. The assumptions are that the data is smooth, i.e. that data that is part of the same classification doesn't change character too rapidly. That data can in some sense be clustered, and that even complex data can be described relatively simply. The semi-supervised method used here is different from the usual methods. First, it uses a single class label, second, the "decision function" is biased by a global parametric model of the data, and third the local similarity functions are compact kernels. The resulting decision function is applied to all the data and the class label is applied to any data that the function evaluates as being part of the same class. This process allows data that is part of a structure of interest to be segmented from, or extracted from, all other data in a data set.

1.5 Multi-Objective Optimization Algorithms

Fitting data to a model where the various objectives conflict or interact is known as multi-objective optimization. The multi-objective optimization problem for fitting the SGM weights the various parts of the model allowing them to be adjusted or traded off relative to one another. As a result a whole sets of "best" solutions

are described as the weights are varied. These optimal solutions are said to be Pareto Optimal [114]. It is up to the modeler to determine when a given solution is acceptable. The usual nonlinear least-squares solver (trust region reflexive in Matlab) is extended to solving a nonlinear multi-objective optimization problem whose objective functions make use of L_1, L_2, and L_∞ norms.

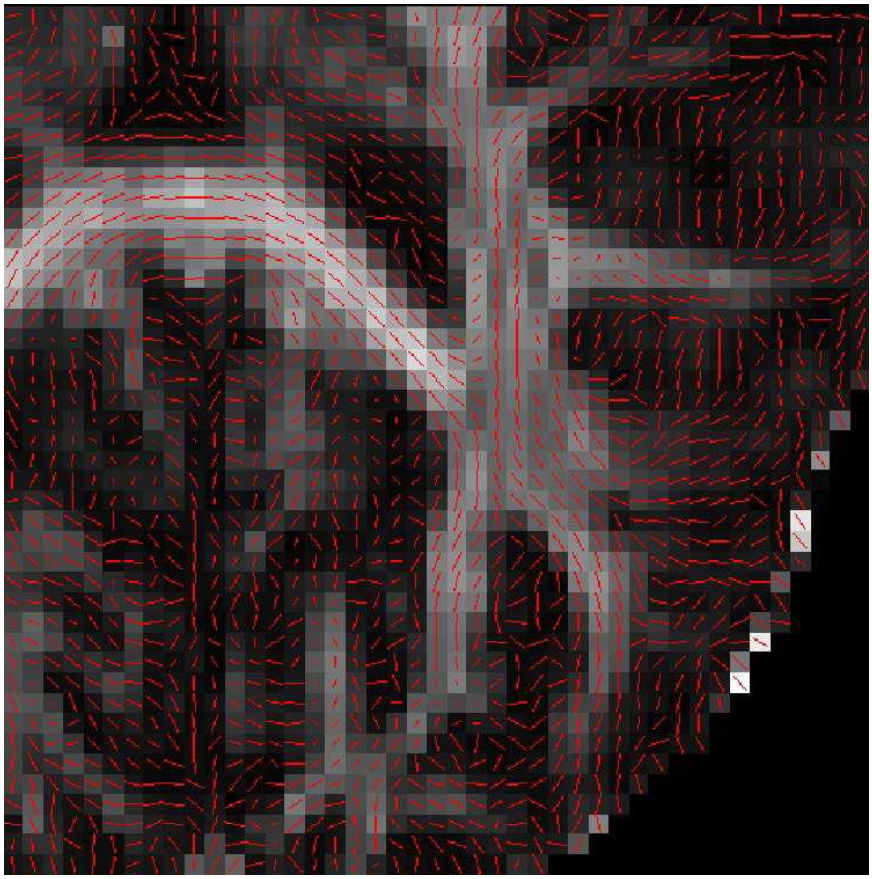

Figure 1.5: Diffusion Weighted Image – axial view slice through a region called the splenium of the corpus callosum. The shading is calculated from the imaging data and indicates the distribution of white matter versus gray matter. The best available resolution is almost 3 orders of magnitude larger than the size of a single myelinated axon. At best, the imaging data is a noisy approximation of the organization of *groups* of myelinated axons.

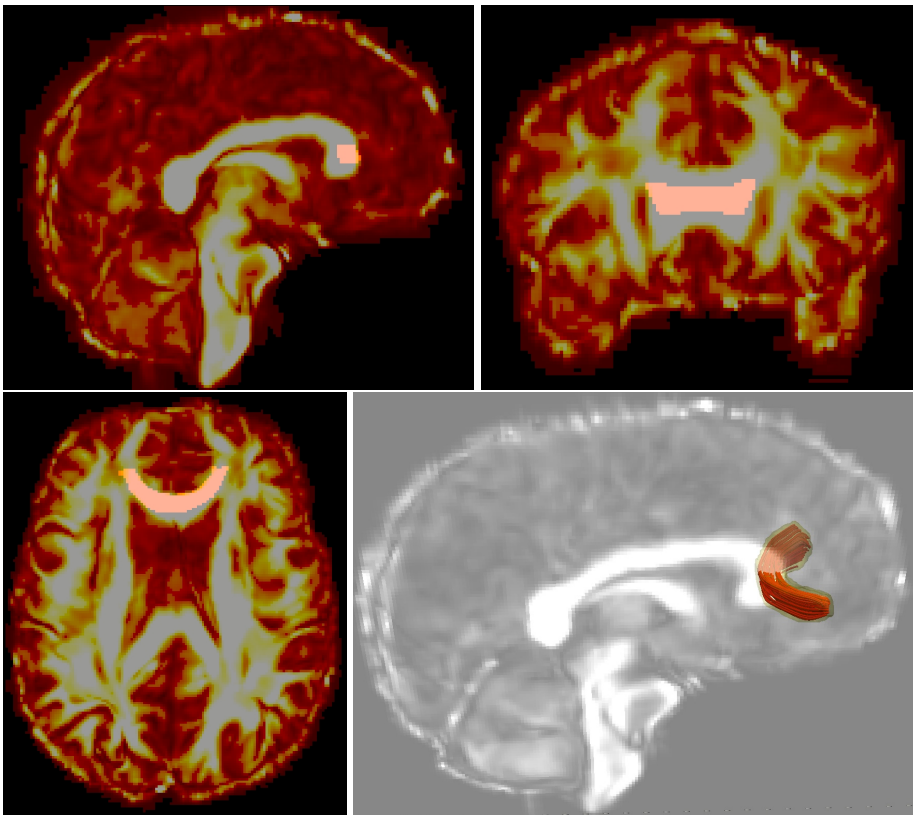

Figure 1.6: Clockwise from lower-left to upper-right. Axial, sagittal, and coronal views with an SGM volume of the genu of the corpus callosum that connects two designated ROIs overlaid. The lower-right image shows manifold curves generated by the model representing the organization of the white matter fibers in the SGM volume.

1.6 Contributions

The elements of an SGM are (1) a novel semi-supervised, semiparametric slab support vector machine, (2) a new geometric entity called a Geometric Median Surface, (3) a global implicit polynomial method for modeling a vector field and by extension a second order tensor field, and (4) an algorithm that fits the SGM model to the data by solving a nonlinear, multi-objective optimization problem whose objective functions make use of L_1, L_2, and L_∞ norms.

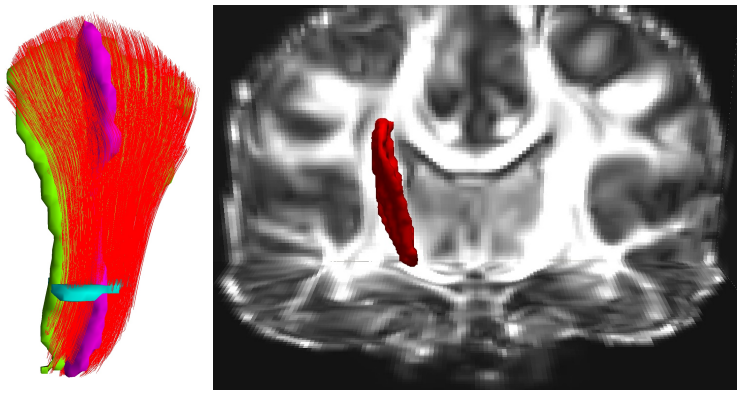

Figure 1.7: (left) is an example of curves modeling a portion of a white matter structure called the corticospinal tract. (right) is a volume of the same structure shown in the position where it is located in the brain.

The SGM is a new approach to discriminating differences in connectivity among different brains, and presents a new framework to further explore how differences in connectivity modulate differences in dynamic processes in the brain. Semiparametric Geometric Modeling is a natural foundation for partitioning, modeling and analyzing the local, global, and volumetric properties of white matter volumetric structures. Unlike all current methods, the SGM inculcates the full structural information available from the data and from assumptions based on the global organization of white matter in the brain into a nonlinear manifold embedded in $\mathbb{R}^3$. This geometric model can be queried to extract geometric and scalar invariants based on the geometry of the white matter rather than on the imaging organization of the data's raw 3D rectilinear grid. The model can partition white matter structures into well defined geometric volumes, surfaces, curves, and points. The major contributions of this dissertation are:

1. A novel method for embedding a nonlinear manifold defined by a second order tensor field in R^3.

2. A novel method for modeling the foliations of a manifold defined by a vector field by a single implicit polynomial function that represents a family

of implicit surfaces. This problem has previously been formulated as a local dynamic problem and solved by PDEs. That method suffers from sensitivity to noise and ignores global structure. Instead, the problem is formulated here as a static global problem that is solved by fitting the gradients of a function with global basis functions and "localizing" the function using the Geometric Median Surface.

3. A novel robust depth function called a Geometric Median Surface (GMS). The GMS generalizes the geometric median point to parametric surfaces in vector fields.

4. A semi-supervised Support Vector method for clustering data on manifolds. This is an extension of the Slab SVM to a single class classifier. It clusters in a manner similar to single-link hierarchical clustering. Data is iteratively added to a labeled class based on compact kernel functions.

5. The SGM manifold model whose local structure is determined by a semi-supervised support vector model and whose global volumetric structure is a function of global, intersecting families of implicit surfaces. For example, in $\mathbb{R}^3$ for a manifold determined by a second-order tensor, the global volumetric structure amounts to the intersections of three families of orthogonally-intersecting implicit functions.

6. A structural query method to organize data on an SGM manifold for statistical analysis. For example Functional Data Analysis to analyze fractional anisotropy (FA) data along white matter fibers.

7. A model that allows analytic calculation of local torsion, curvature and other differential geometric quantities at any point in the volume modeled by an SGM.

8. A model that allows for rigid body transformations, for example rotations and translations, of an extracted volumetric structure by linear transforms of the coefficients of the implicit functions that are the output of the SGM.

9. A new definition of white matter connectivity is defined based on the SGM's ability to partition white matter structures into well defined volumes.

1.7 Outline of the Thesis

The remainder of this dissertation is organized as follows.

Chapter 2 presents an overview of the basic organization of neural tissue in the brain. Focused primarily on the properties of white matter tracts.

Chapter 3 provides an overview of Diffusion Weighted Magnetic Resonance Imaging (DWI) and what it measures in white matter. The balance of Chapter 3 looks at current methods of modeling and analyzing white matter structures.

Chapter 4 reviews the fundamental concepts and constructs that underlie the modeling method that is the subject of this thesis. The most important abstract elements are the definition of manifolds, foliations, implicit algebraic surfaces as foliations of a manifold, and total ordering of foliations. The most important algorithmic notions reviewed are semiparametric models, manifold learning, and semi-supervised learning. As these ideas are presented some of their relationships to the application domain are briefly described.

Chapter 5 presents a new modeling framework called Semiparametric Geometric Modeling (SGM). An SGM brings together elements of Differential Geometry, Differential Topology and Machine Learning to model the data at multiple scales on multiple structures. The resulting algorithm enables automatically extracting and modeling white matter structures from DTI data.

Chapter 6 presents the algorithm for fitting an SGM. Finding the "best" fit to the data of each part of the model and then combining the parts would very likely give a model that is a poor fit to the data. Since there are multiple parts or objectives in the model, a better approach is to combine the objectives together to form a multi-objective function. The multi-objective formulation and the algorithm to fit it is described.

Chapter 7 introduces methods for querying geometric properties of the SGM.

It begins by demonstrating how to extract submanifolds – volumes, surfaces, curves, and points from the model. Next, it defines how to map points from image space to manifold space and from manifold space to image space. This chapter also specifies how to sample submanifolds so as to organize the data for Functional Data Analysis, and how to use the SGM to calculate invariants at each point on the manifold. Next, concepts behind rigid body transformations of the manifold model are presented. This chapter also includes the definition of connectivity used in the SGM. Finally, orientation statistics are presented as a method to access how well the SGM fits the data and as a method to detect small-scale structures that might not be seen by the global model.

Chapter 8 presents an implementation of SGM and uses it to analyze a Diffusion MRI data set and is able to detect, and localize group differences in two white matter structures that have been implicated in Autism Spectrum Disorder. The software that implements the SGM includes software to (1) build SGM models, (2) implements sampling and mapping functions, (3) outputs SGM volumes at voxel (NIFTI format) and subvoxel levels (Matlab mat format), (4) outputs SGM curves, $\mathcal{C}_{\parallel}, \mathcal{C}_{\perp}$, and $\mathcal{C}_{\Vdash}$ in Camino, Trackvis formats, and Matlab mat format (5) maps users choice of invariants such as FA onto SGM curves and outputs each curve as a spline function ready for Functional Data Analysis, and (6) provides a wrapper around third party FDA package to do the analysis and generate reports.

Chapter 9 reviews the contributions of this thesis and outlines possible future work. Additionally, new methods of solutions are suggested. Also, ways of extending these methods to other imaging modalities and the incorporating clinical data are discussed.

Chapter 2

Organization of White Matter of the Brain

2.1 Introduction

In this chapter we briefly review the basic organization of neural tissue in the brain. We focus primarily on white matter volumetric structure and the fiber bundles that make it up. Identifying these white matter regions, characterizing their volumetric structure, and determining what gray matter regions of interest (ROI) they connect, is important for understanding both normal brain function and disease processes.

2.2 Volumetric Structure and Organization of Neural Tissue of the Brain

Early anatomical studies of the brain revealed two major types of tissue: white matter and gray matter as shown in Figure 2.2. Looking more closely, histological studies revealed that what appeared to be two separate types of tissue shared a common component—the neuron (See Figure 2.1).

The part of the neural tissue associated with gray matter is predominantly unmyelinated axons and cell bodies. The cell bodies may be considered processing nodes with unmyelinated axons as the short range local connections between neural cell bodies. Unmyelinated axons can be packed very tightly allowing many short range connections but have comparatively low signaling velocity. Conversely, myelinated axons in white matter are bulkier but have additional structure that allows much higher signaling velocities.

Myelinated axons are axons wrapped in myelin sheaths. The myelin is produced by oligodendrocytes (Figure 2.1). A single oligodendrocyte may myelinate as many 50 axons, in effect helping to define and constrain the local geometry of groups or bundles of axons. This suggests that the myelinated axons are unlikely to be arranged in a tangled mass but have a more coherent local organization. Bundles of myelinated axons are called white matter structures or fiber tracts. The fibrous organization of the myelinated axons is apparent in the dissection of the corticospinal tract shown in Figure 2.2.

Until recently the fibrous organization has defined the primary geometric consideration in modeling the organization of white matter structures. In this case only the geometry parallel to the bundles is taken into account. The reason for this is that the data sampling method, Diffusion Magnetic Resonance Imaging (DMI) relies on the freedom of water to diffuse along the myelin parallel to the underlying axon bundles. We make use of empirical evidence that the sampling method may also be used to infer the geometry orthogonal to, or across the orientation of, the axon bundles. For example, modeling sheet-like, tube-like or more general kinds of organization. We consider the fiber bundle and additional geometric organization orthogonal to the fiber bundle locally to be the mesoscale geometry of the brain. The macroscale or global organization of the main body of the brain has yet to be fully characterized but has been proposed to be tied to the embryonic axes of development.

The general body plan of all chordates is known to follow a set of intrinsic geometric axes known as embryonic axes of development [52]. These are, the

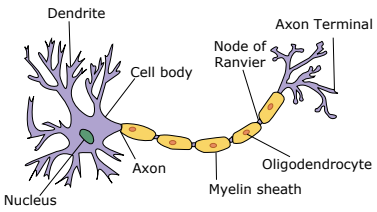

Figure 2.1: Schematic of a neuron. The cell body containing the nucleus and the dendrites and axon terminal comprise the gray matter of the brain. The myelin sheathed axons are the long range connections of the brain and make up the white matter fibers of the brain.

rostro-caudal, the medio-lateral (or proximo-distal), and the dorso-ventral axes. The brainstem and midbrain have also been shown to have structural organization that follows these axes. The work in this thesis concentrates predominantly on the mesoscale organization of the brain but we return to the macroscale organization in Chapter 9.

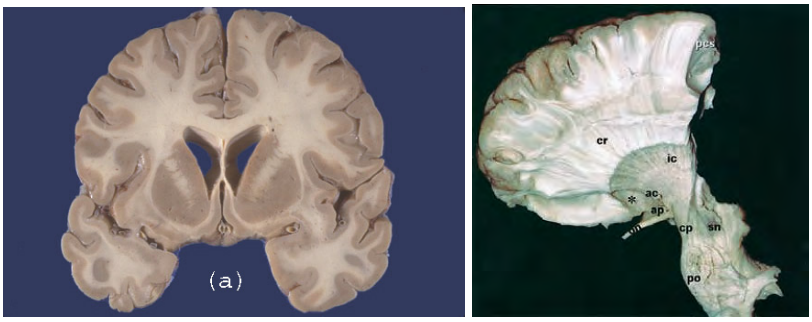

Figure 2.2: Dissection of the corticospinal tract. (left) Coronal view slice through the brain demonstrating the gray matter and white matter structures at the gross anatomical level. (right) Brain with portions of the white and gray matter removed showing the midbrain to the cortex. The fibrous organization of white matter can be seen. From [173].

2.3 Discussion

This chapter reviewed a few properties of white matter fibers and fiber bundles. It is the modeling of the large scale organization of these white matter bundles that is the object of this thesis. Clearly, before the organization these substructures can be modeled it is necessary to acquire information about it's structure *in vivo*. An imaging modality that provides the information that is needed is called Diffusion Weighted Magnetic Resonance Imaging. The next chapter, Chapter 3, reviews the principles of Diffusion Weighted Magnetic Resonance Imaging and some of the current methods for modeling white matter organization.

Chapter 3

Diffusion Weighted Magnetic Resonance Imaging

This chapter provides and overview of Diffusion Weighted Magnetic Magnetic Resonance Imaging (DWI) and what it indicates about the local organization of white matter. We then look specifically at Diffusion Tensor Imaging (DTI) which has historically been the major imaging method for analyzing white matter tissue. The balance of Chapter 3 looks at current methods of modeling and analyzing white matter structures.

A variety of methods are available to study white matter tissue. These range from histology and chemical tracers to DWI. Histology and tracers are excellent at the very small scale but are impractical for large numbers of neurons over long distances and are impractical for *in vivo* studies in humans. DWI on the other hand is ideal for studying white matter at a large scale and in living human beings.

DWI samples the water diffusivity properties of tissue that correlate with the microstructure and organization of the tissue. However, the best available resolution is almost 3 orders of magnitude larger than the size of a single myelinated axon. As a result, what is measured are the local properties of *bundles* of myelinated axons. Fortunately, DWI is very sensitive to the organization and microstructure of these bundles of myelinated axons [59] [18] [159].

A single DWI voxel may be a sample of a single tissue type or substance. For example, white matter, gray matter, or Cerebrospinal Fluid (CSF). Alternatively, a single voxel may contain a combination of these tissue/substance types. Additionally, a voxel may be the result of sampling an area where axon bundles cross, split or merge. Instances where a voxel is the result of sampling multiple tissue/substance types or where the fibers have a complex local geometry are known as the *partial volume problem*.

There are a number of variations of DWI that may be used to sample white matter, but by far the most common is Diffusion Tensor Imaging (DTI) and is the imaging modality we use in our model. See [81] for a detailed examination and survey of DWI methods.

3.1 Diffusion Tensor Imaging

The **Diffusion Tensor** is represented by a positive definite 3x3 matrix D with 6 degrees of freedom. The Stejskal and Tanner [81] equation:

$$S_\iota = S_0 \epsilon^{b_\iota \otimes \mathbf{D}} \tag{3.1}$$

is the basis for the estimating the **Diffusion Tensor** [81], where **b** is the matrix that encodes measurement parameters, **D** is the Diffusion Tensor, S_0 is the image acquired with no diffusion encoding gradients, and $S_\imath$ is the image acquired with each change of the diffusion gradient. Note that the diffusion tensor model produces a highly smoothed orientation distribution making it difficult to infer how the tissue orientation might vary spatially within a sampled voxel.

This is especially challenging in voxels where partial voluming occurs. DTI alone cannot provide enough information to deal effectively with the partial volume problem. The Diffusion Tensor **D** in Equation 3.1 is represented by a symmetric 3x3 positive definite matrix. The eigen decomposition of **D** results in 3 eigenvectors and their associated eigenvalues, $\vec{\varepsilon}_\imath, \lambda_\imath \; \imath = 1,2,3$. Sorted by largest to smallest eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$. Following [193] we relate the geometry

of **D** as follows. The primary diffusion direction or vector is interpreted to be aligned parallel with the axis of the underlying fiber bundle. Specifically, λ_1 is the eigenvalue and $\vec{\varepsilon}_{\parallel} = \vec{\varepsilon}_1$ is the associated eigenvector parallel to the fiber bundle. For the secondary diffusion direction, which we consider the dispersion of the fiber bundle, λ_2 is the eigenvalue and its associated eigenvector $\vec{\varepsilon}_{\perp} = \vec{\varepsilon}_2$ is the direction of the dispersion. Finally, the tertiary diffusion direction is the normal to the first two defined by λ_3 and its associated eigenvector $\vec{\varepsilon}_{\shortmid\!\vdash} = \vec{\varepsilon}_3$.

A schematic of a Diffusion Tensor is shown in Figure 3.1. Our method assumes that for any small area on a given fiber bundle the tensors in neighboring voxels, each with a given $\vec{\varepsilon}_{\parallel,\mathrm{j}}$, $\vec{\varepsilon}_{\perp,\mathrm{j}}$, and $\vec{\varepsilon}_{\shortmid\!\vdash,\mathrm{j}}$ for the voxels $\mathrm{j} = 1...\mathrm{N}$ with N being the number of voxels, are approximately parallel and vary locally about a common orientation. However, large scale changes in the organization of the tissue, for example bifurcation points, do occur where tensors change orientation rapidly and we will describe these exceptions in Chapter 8.

The diffusibility properties measured by DTI that correlate with tissue organization are important when trying to estimate tissue integrity and composition. Scalar values that are independent of local tissue orientation are particularly useful for analysis. A number of invariants derived from the Diffusion Tensor, **D**, have been shown to be correlated with characteristics of diffusivity in white matter. Beginning with a definition of Mean Diffusivity, **MD**, as

$$\mathbf{MD} = \mathbf{Tr}(\mathbf{D})/3 \tag{3.2}$$

where $\mathbf{Tr}(\cdot)$ is the trace of the Diffusion Tensor. The trace is a well known tensor invariant. Letting $\hat{\lambda} = \mathbf{MD}$, Fractional Anisotropy **FA** [13] is defined as

$$\mathbf{FA} = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1-\hat{\lambda})^2+(\lambda_2-\hat{\lambda})^2+(\lambda_3-\hat{\lambda})^2}{\sqrt{\lambda_1^2+\lambda_2^2+\lambda_3^2}}}, \tag{3.3}$$

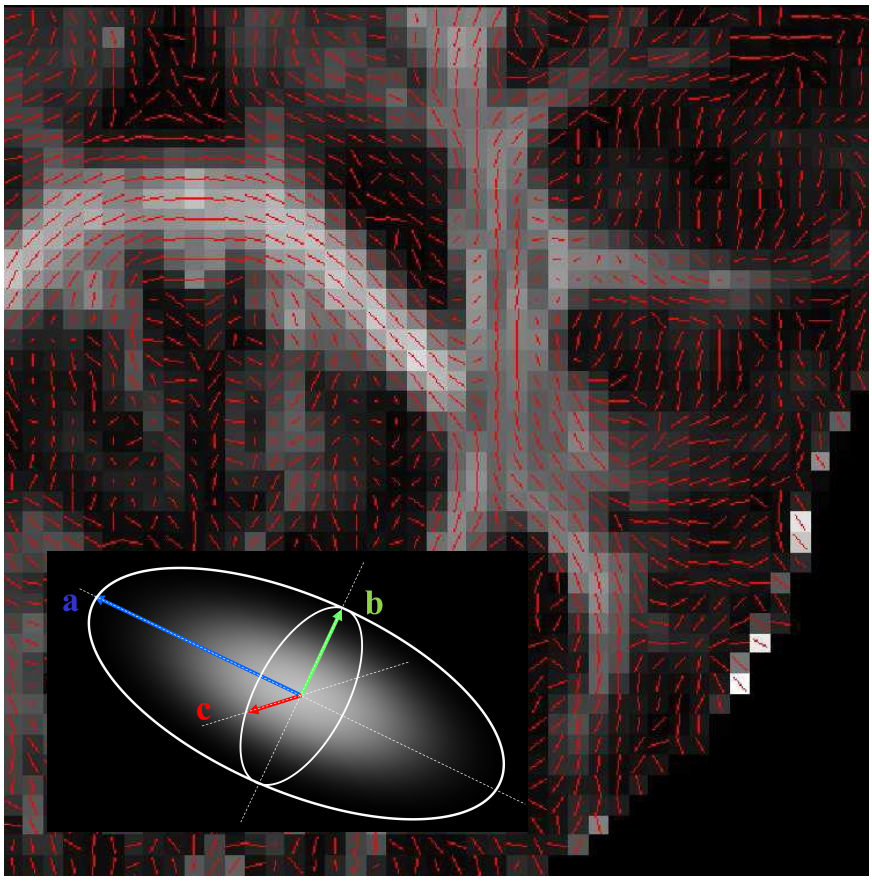

Figure 3.1: Diffusion Tensor: a) principal diffusion direction $\vec{\varepsilon}_{\parallel}$. b) Normal diffusion direction $\vec{\varepsilon}_{\perp}$. c) Binormal diffusion direction $\vec{\varepsilon}_{\shortmid\!\perp}$. The principal diffusion direction is aligned with white matter structures.

and Relative Anisotropy **RA** [13] as

$$\mathbf{RA} = \sqrt{\frac{1}{3}} \frac{\sqrt{(\lambda_1 - \hat{\lambda})^2 + (\lambda_2 - \hat{\lambda})^2 + (\lambda_3 - \hat{\lambda})^2}}{\mathrm{Tr}(\lambda)}. \tag{3.4}$$

MD, **FA**, and **RA** are the scalar data that are analyzed using sampling methods from our model in Chapter 8.

In Section 3.2 we review current methods used to characterize white matter

fiber bundles. One example of tracing fiber bundles by streamlines is shown in Figure 3.2.

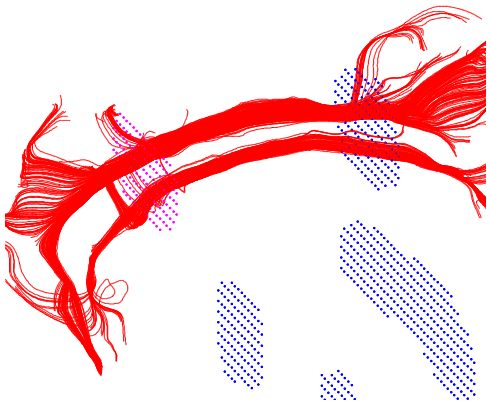

Figure 3.2: Streamline tractography: The streamlines in red trace along the fibers of the left and right cingulum bundles.

3.2 Deterministic Tractography

The term **tractography** has a number of meanings. Unless noted otherwise we use it here to refer to the deterministic streamline algorithms used to analyze white matter structure. The space curves that result from applying these models will be referred to as **streamlines**. The term **tract** refers to a physical white matter structure. Finally, the term **fiber** refers to bundles of myelinated axons.

Identifying which bundles of myelinated axons belong to a given white matter tract, characterizing their structure, and determining what gray matter ROIs they connect, is important for understanding both normal brain function and disease processes.

Tractography can be thought of as a set of methods for finding an unknown nonparametric function whose solution defines a space curve through the data that corresponds to the axis of a white matter fiber bundle. This is accomplished

by propagating a streamline from an arbitrary starting or seed point through a vector or tensor field defined by the diffusion data.

While this has been a very fruitful approach, it has a number limitations. [14] and [94] have shown that the farther a streamline is propagated from a seed point, the greater the uncertainty that that tract represents a real connection between ROIs. In addition, tracts are 1D objects that do not interact with each other. This means that an arbitrary number of tracts may cross a single voxel. In fact very little can be inferred about the geometry between two nearby streamlines. Other than some local smoothness constraints on the propagation methods, in general there are no global constraints. An exception to this are constraints that require streamlines to pass through or avoid ROIs in the data [79] and the requirement that each streamline is constrained to lie on a parametric curve [1].

Questions that anatomists might like to ask tend to be global questions about membership and strength of connection [82]. There are significant challenges in using streamline tractography to address these questions. For example, deciding which streamlines are part of a specific white matter structure using the intersection constraint mentioned above is challenging. Due to noise, some streamlines that should connect two ROIs may not do so. Conversely, again because of noise, some streamlines that should not intersect the ROIs, appear to do so.

Using streamlines to determine the strength of connection between two ROIs is also challenging. The most common method is to count the number of streamlines that intersect the ROIs. This is, in effect, an effort to estimate the volume of the white matter connecting the two ROIs. In addition to the noise related issues mentioned above, there is the question of over-counting. That is, arbitrarily many streamlines may pass through a voxel, and the greater the distance between the ROIs, the more likely streamlines are to intersect a given voxel. This results in an overestimation of the volume of tissue connecting two ROIs. The current method for correcting for this is *ad hoc* and not entirely satisfactory. Some form of rejection sampling [46] might be a better choice for estimating this volume using streamlines.

In contrast, our method provides well-defined definitions of both membership and volume.

3.2.1 Deterministic Streamline Tractography

Historically, deterministic streamline tractography was borrowed from methods used to visualize the organization of vector fields [28] [154]. Initially it was used as a means to visualize and group white matter tensors based on the principal diffusion direction ($\vec{\varepsilon}_{\|}$) vector. Subsequently, streamlines were extended to quantitative analysis. We begin by looking at streamlines as solutions to families of Ordinary Differential Equations (ODEs) on vector fields of $\vec{\varepsilon}_{\|}$. Then we look at methods that make use of the full Diffusion Tensor to find streamlines that are solutions to Partial Differential Equations (PDEs).

First, we consider two properties that are important to both methods: Smoothness and ambiguity of the "direction" of the vectors derived from the diffusion tensor. Given that we are tracing physical structures with known geometric properties that by inspection can be considered Lipschitz continuous [147], smoothness in this case amounts to choosing a local cone as shown in the white area in Figure 3.3. If the apex of the cone is slid along the curve and no other nearby point intersects the body of the cone, it is considered Lipschitz smooth.

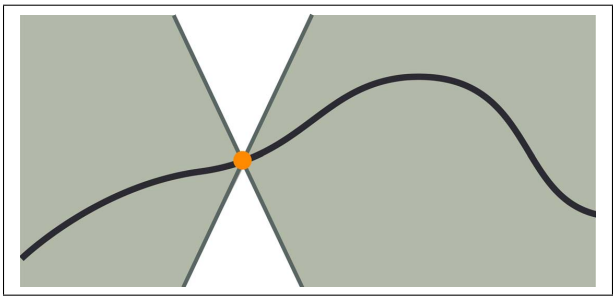

Figure 3.3: Lipschitz continuous means that the rate of change in a continuous function is bounded. For example, in this sketch if the curve changes so rapidly that it enters the white region it would not be considered Lipschitz continuous.

It is, however, useful to define a constant $\mathbf{L}$, such that

$$\|f(x_1) - f(x_2)\| < \mathbf{L}\|x_1 - x_2\|, \tag{3.5}$$

where $\mathbf{L}$ specifies the maximum local rate of change in the function f. Deterministic streamline techniques choose a Lipschitz constant as a hard stopping constraint when generating streamlines. We return to Lipschitz continuity when discussing our methods.

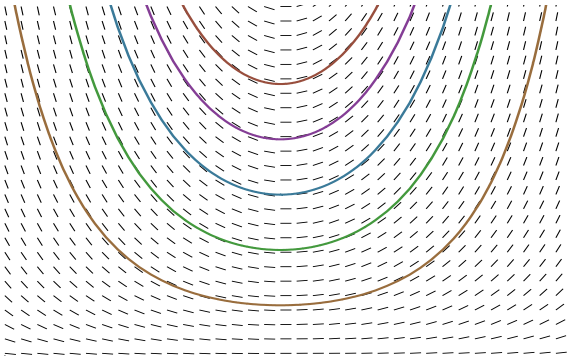

Figure 3.4: Slope or orientation fields graphically represent the solution set of first-order differential equations (ODE). Streamline tractography traces pathways through slope fields by propagating along the solution set of ODEs.

The diffusion properties estimated using DTI give us a $\vec{\varepsilon}_{\|}$ that can't be said to have a "direction." That means there is really no preference for $\vec{\varepsilon}_{\|}$ versus $-\vec{\varepsilon}_{\|}$. We therefore consider the field to be a slope or tangent field that can be solved using families of ODEs. Given a set of seed points or initial values the ODEs may be solved iteratively by numerical methods as we will see in the next section.

3.2.2 Generating Streamlines using ODEs

$\mathbf{V}(\tilde{\mathbf{x}})$ may be rewritten as $\mathbf{V}([\mathbf{x_1}(t), \mathbf{x_2}(t), \mathbf{x}_{\ldots\mathbf{n}}(t)]$, in $\mathbb{R}^n$. Then,

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{V}_i(\tilde{\mathbf{x}}), t > 0 \tag{3.6}$$

is an autonomous linear ODE where t is a parameter of the integration and can be thought of as local distance along the length of the curve $\mathbf{x}_i$ where $i \in 1...N$ is a family of curves, $\tilde{\mathbf{x}}(\mathbf{t})$, that are tangent to the vectors $\mathbf{V}_i(\tilde{\mathbf{x}})$.

The most common methods for numerically solving these ODEs are local applications of Euler's method or 4^{th} order Runge-Kutta. See [12, 85] and a survey [205] for Runge-Kutta methods applied to streamline tractography. A step size is chosen, and local integration is applied to estimate the next next point to be included in the function. It is common to choose a Lipschitz constant, **L**, (Figure 3.3) such that a large abrupt change in direction in the streamline causes the ODE to halt. Another stopping criterion is, when the value of **FA** in the voxel concurrent with the local solution is less than a constant.

3.2.3 PDE–Level Set Propagation

A vector field in $\mathbb{R}^3$ can be thought of as a set of local systems of partial differential equations (PDEs):

$$V(\vec{x})_i = \left[\frac{\partial F}{\partial \vec{x}}\right]. \tag{3.7}$$

The solutions of this kind of PDE turn out to be surfaces that are the level sets of a function. In order to find unique solutions, some further constraints must be applied.

Level set propagation methods use all of the Diffusion Tensor information at a voxel to do tractography. [205] is a survey of different tractography techniques including descriptions of wavefront propagation. [124] and [30] used evolving level set surfaces and a variation of directed Fast Marching to determine a maximum diffusion direction. [69] used Navier-Stokes fluid flow simulation followed by an Active Contour method [200] to generate streamlines. Tensor Deflection [96] is a wavefront algorithm that uses a weighting scheme rather than velocity-based methods or pseudo viscosities. All these techniques apply criteria that allow for a unique solution that determines the direction of propagation for the next step in tracing the streamline.

3.3 Volumetric Methods

While streamline tractography has some useful properties, there is a need to be able to analyze white matter volume as well. As described in Section 3.2, questions about strength of connection between two ROIs are closely related to the volume of tissue connecting those ROIs. One approach to attempt to characterize volumes using streamlines, is clustering.

3.3.1 Tract Clustering

Streamline tractography produces 1D curves that then need to be agglomerated to represent physical white matter structures. One method to accomplish this it to define ROIs and choose those tracts that pass through them and obey certain rules. For example, tracts may be chosen such that they pass through ROI A and ROI B but not ROI C. More sophisticated clustering methods include spectral methods [83, 125, 189, 210]. In this case features are extracted from the curves and spectral methods are used to find clusters. Probabilistic clustering methods [55, 105] cluster tracts by finding a set of minimum cost paths that connect two given regions of a brain. The minimum cost is chosen and paths that fall within some ϵ cost of that path are included. While this method does a good job of localizing a set of tracts to a volume there is no real geometric relationship between the tracts at any point besides the end points.

3.4 Volumetric Methods without Streamlines

Since our method does not use streamlines we next describe other volumetric methods that also do not require streamlines.

3.4.1 Tube Fitting

[64] models white matter bundles using a tube fitting algorithm based on Principal Curves. Guided by $\vec{\varepsilon}_{\parallel}$, their algorithm starts at a seed point. An implicit surface orthogonal to $\vec{\varepsilon}_{\parallel}$ at the seed point is determined. The surface is constructed so as to be the level set of a bivariate normal function (an ellipse). Nearby points that project onto the surface are selected and added to the tube. A medial point in the current surface is determined, and then a step is taken, a new level surface is then constructed, and new nearby points are added. The set of medial points are connected and become the medial axis of the tube. As a final step the boundary of the tube is estimated. The result is a tube parameterized by a medial axis. As noted in [64], major cortico-cortico fiber bundles would be poorly modeled by a tubular structure. However, structures such as the Cortico Spinal tract might not be well modeled by such a tube-like structure. Like other methods, it fails to take into account the full information available from Diffusion Imaging. While taking volume information into account, this method doesn't allow bifurcations of the extracted structure.

3.4.2 Deformable Models

Deformable models have a conceptual relationship with our method. M-reps [130, 203] and other deformable models [112] fit various geometric and algebraic objects to data sets locally and iteratively, unlike our method, that fits data sets globally. Recent work [23] has applied this approach specifically to diffusion imaging [54, 158]. Most deformable models fit the overall shape of the structure. They then impose an arbitrary coordinate system unrelated to the internal structure to analyze the interior structure.

3.5 Discussion

This chapter reviewed current methods of modeling white matter structures as well as some of the mathematical concepts behind them. The next chapter begins the process of putting together the mathematical concepts and tools that are at the foundation of our new modeling method.

Chapter 4

Mathematical Background

4.1 Introduction

This chapter introduces the fundamental concepts and constructs that underlie the modeling method that is the subject of this thesis. The most important abstract elements are the definition of manifolds, foliations, implicit algebraic surfaces as foliations of a manifold, and total ordering of foliations. The most important algorithmic notions reviewed are semiparametric models, manifold learning, and semi-supervised learning. As these ideas are presented some of their relationships to the application domain are briefly described.

Figure 4.1 is a representation of the white matter organization in a human brain using streamline tractography. The number of streamlines chosen is *ad hoc* and can often be more that 100,000. This is an extremely complex and large data set. In fact, one way of describing the complexity while making minimal assumptions would be to say the degrees of freedom, or dimensionality, of the data are the number of streamlines times some number of points sampled on each streamline. Even if the streamlines themselves are represented parametrically as in [1, 36], the dimensionality is still extremely high. One way to reduce the dimensionality of the data and in some sense make it easier to analyze is to assume that it lies on, or is sampled from, a lower dimensional structure.

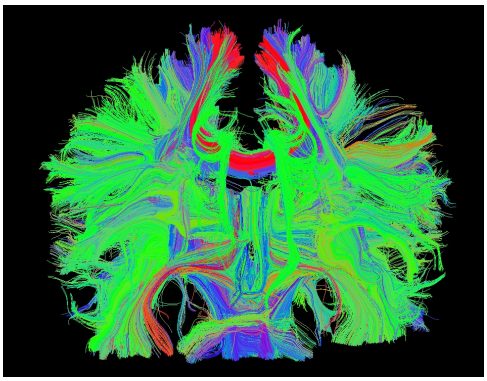

Figure 4.1: Visualization of the organization of the brain using streamline tractography.

4.2 Manifolds

"Think Globally, Fit Locally" [141] is a good description of geometric manifold modeling. In order to be a geometric manifold a solution space must have certain properties. First, it must be a metric space. A metric space $\mathbb{M}$ must have a suitable local distance function or norm with the follow properties:

1. $d(x,y) \geqslant 0$ for all $x,y \in \mathbb{M}$
2. $d(x,y) = 0$ if and only if $x = y$
3. $d(x,y) = d(y,x)$ for all $x,y \in \mathbb{M}$
4. $d(x,y) \leqslant d(x,z) + d(y,z)$ for all $x,y,z \in \mathbb{M}$ (triangle inequality)

Let $(\mathbb{M}, d)$ be a metric space. For each point x and $\epsilon > 0$, the norm is confined to an $\epsilon - \mathtt{ball}$ or $\epsilon - \mathtt{neighborhood}$ about each point x.

$$D(x,\epsilon) = y \in \mathbb{M} | d(x,y) < \epsilon, \tag{4.1}$$

where $D(x,\epsilon)$ are open subsets about each x. We are working in a space of continuous functions and topological spaces that are the natural domain of such

functions [29, 195]. A topological space is defined as a set X of points with open subsets O that satisfy the following:

1. The empty set, $\emptyset \in O$
2. $X \in O$
3. $I = O_i \cap O_j \; \{i, j\} \in \{1...k, k < \infty\}$: $I \in O$
4. $J = O_i \cup O_j \; \{i, j\} \in \{1...k, k \leqslant \infty\}$: $J \in O$

We need two more definitions on metric spaces:

Definition 4.1. ***Homeomorphism***

1. *Homeomorphism – a continuous function between two topological spaces that is bicontinuous.*
2. *A space is a locally Euclidean space* X *of dimension* N *if* $\forall x \in X$, $\exists \, D(x, \epsilon) \subset X$ *and* $D'(x, \epsilon) \subset \mathbb{R}^N$ *and a homeomorphism* $\phi : D(x, \epsilon) \mapsto D'(x, \epsilon)$.

Definition 4.2. ***Manifold***

A ***manifold*** $\mathfrak{M}$ *of dimension* N *is a topological space with the following properties:*

1. $\mathfrak{M}$ *is Hausdorff*
2. $\mathfrak{M}$ *is locally Euclidean*
3. $\mathfrak{M}$ *has a countable basis of open sets*

Finally, it is assumed that manifolds are compact but not necessarily bounded.

When a manifold is not compact it can, in general, be made compact using a process called compactification.

Definition 4.3. ***Compactification***

Compactification *or making a manifold compact is accomplished by adding points at infinity. This means transforming the problem into the real projective space* P^3. *The resulting manifold is both compact and closed.*

We rarely explicitly perform compactification. However, when we need these properties it is assumed that we can compactify the manifold in order to perform whatever analysis or operations are required, and then convert the result to back to $\mathbb{R}^3$.

4.3 Foliations

Informally, foliations are non-intersecting surfaces that partition a manifold into connected subsets of points [116, 170]. Additionally, they may be thought of as filling the space of a manifold. Another way of stating this is that a *foliation* is a decomposition of a manifold into a *union* of locally parallel submanifolds of lower dimensionality. For example, the rings of a tree trunk may be considered a foliation of a manifold represented by a generalized cylinder. The layers of an onion are a foliation of a sphere. Or the simplest example of a foliation is the pages of a book. Following [201] we define foliations as:

Definition 4.4. ***Foliations and leaves***
Foliations and leaves *[138] Let $\mathfrak{M}^N$ be a manifold and let $\mathfrak{F} = \{\mathfrak{L}_\alpha\}$ denote a partition of $\mathfrak{M}^N$ into a family of B disjoint pathwise-connected subsets. Then $\mathfrak{F}$ is a foliation of codimension c, with $0 < c < N$, if there exists a cover of $\mathfrak{M}^N$ by open sets U, each equipped with a homeomorphism $h : U \mapsto \mathbb{R}^N$ or $h : U \mapsto \mathbb{R}^N_+$ which sends each nonempty component of $\mathfrak{F}_\alpha \cap U$ onto $\mathbb{R}^{N-c} \times \alpha$ in $\mathbb{R}^{N-c} \times \mathbb{R}^c = \mathbb{R}^N$. $\mathfrak{L}_\alpha$ is called a leaf.*

1. $\bigcup_{\alpha \in B} \mathfrak{L}_\alpha = \mathfrak{M}$
2. *For every $\alpha, \beta \in B$ if $\alpha \neq \beta$ then $\mathfrak{L}_\alpha \bigcap \mathfrak{L}_\beta = \emptyset$*
3. *For any point $p \in \mathfrak{M}$ there exists a local chart $(U, \phi) \in A$ where A is an atlas, $p \in U$ so that if $U \bigcap \mathfrak{L}_\alpha \neq \emptyset$ for some $\alpha \in B$ the components of $\phi(U \bigcap \mathfrak{L}_\alpha)$ are subsets of parallel affine planes*

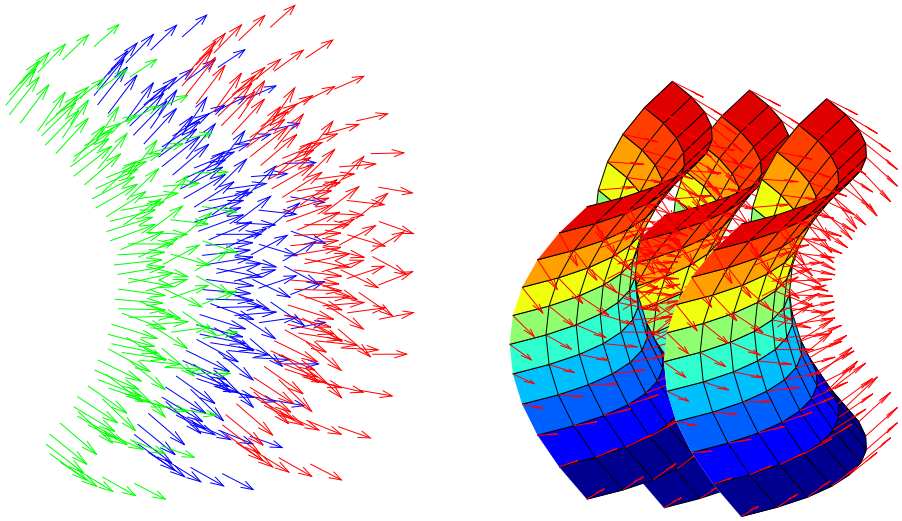

Figure 4.2: (left) Shows a vector field in space. (right) Vector field with level surfaces. When the vector field describes a manifold, the level surfaces correspond to foliations of the manifold.

A vector field V on a smooth manifold $\mathfrak{M}$ gives rise to a decomposition of $\mathfrak{M}$ by the integral curves of V [170]. Integral curves are the solutions of ODEs and represent co-dimension 2 foliations. On the other hand, the solution surfaces of PDEs [120] are codimension 1 foliations.

There is an important distinction between the two types of foliations. Codimension 2 foliations in general have no partial order on $\mathfrak{M}^3$. Conversely, [123] has shown partial ordering for codimension 1 foliations.

Definition 4.5. ***Partial ordering of foliations:*** $\bigcup_{\alpha\in B} \mathfrak{L}_\alpha$ *is the set of all leaves of* $\mathfrak{M}$. *Consider the relation* $\leqslant$ *on the set of leaves. We say* $\mathfrak{L}_1 \leqslant \mathfrak{L}_2$ *if and only if* $\mathfrak{L}_1 \subset \overline{\mathfrak{L}}_2$. *Then*

1. $\mathfrak{L}_1 \leqslant \mathfrak{L}_1$
2. *if* $\mathfrak{L}_1 \leqslant \mathfrak{L}_2$ *and* $\mathfrak{L}_2 \leqslant \mathfrak{L}_1$, *then* $\mathfrak{L}_1 = \mathfrak{L}_2$

3. if $\mathfrak{L}_1 \leqslant \mathfrak{L}_2$ *and* $\mathfrak{L}_2 \leqslant \mathfrak{L}_3$*, then* $\mathfrak{L}_1 \leqslant \mathfrak{L}_3$

Level surfaces of families of implicit differentiable functions, $F(x)$, are congruent to co-dimension 1 foliations on $\mathfrak{M}^3$. Implicit functions can be shown to be **totally ordered** with the ordering parameterized by some function $g(t)$ where **t** is an index variable.

In this case we have an implicit function,

$$F(x) = g(t), t \in \mathbf{R}; (-\infty, \infty) \tag{4.2}$$

where $g(t)$ is a signed function and is monotonically decreasing to the left of 0 and monotonically increasing to the right of 0.

4.4 Algebraic Implicit Surfaces

The study of the differential geometry of curves and surfaces has a long history [51]. Implicit algebraic and implicit parametric curves and surfaces have been used extensively in Computer Vision, Computer Graphics, and Medical Imaging, with many applications [11, 22, 113, 144, 165, 208]. With the exception of work related to [165], most methods are nonparametric and applied to closed surfaces. Our method differs from these in that the surfaces are implicit algebraic surfaces congruent to foliations of a manifold. More specifically, we use families of implicit algebraic functions to model the internal structure of volumes rather than a single implicit function to model a surface.

In Chapter 5 we present our model and show that among other properties volumes, curves and points are well defined on the manifold by our model. These are important properties in our context because for example, connectivity between gray matter regions is defined as a function of the volume of white matter connecting those regions.

4.5 Manifold Learning

Classical Manifold Learning algorithms [62] assume a single, possibly nonlinear, low-dimensional manifold embedded in a space of high ambient dimension; in other words, the data are assumed to be sampled (with noise) from *one* underlying manifold. Among the better known methods are Isomap, Locally Linear Embedding, Laplacian Eigenmaps, and Semidefinite Embedding [97]. All of these approaches are nonparametric, assume only one manifold, and require extensions to handle out of sample data, that is, new data. This problem is well studied and various powerful algorithms are known. On the other hand, the generalization of these methods to more complex data sets, specifically those containing multiple manifolds, remains non-trivial and an area of current research. The available tools for solving these kinds of problems are far less developed.

Some Multi-Manifold Learning algorithms begin by partitioning the data using an estimate of local structure. They then apply standard Manifold Learning [183] or Semi-Supervised Learning [61]. One, based on Gestalt ideas [118], uses a voting scheme. [48] and [118] are approaches that do not look for an embedding. They estimate the manifold and then work with data in the original space. Neither estimates a global model. Our model is global and continuous. This allows for proper handling of out-of-sample data naturally while more local methods require searching for neighborhoods that are “close enough" to the model to be included.

4.6 Kernel Machines

Our multi-manifold learning and modeling methods make use of kernel machines. This section gives a brief overview of this very large area of study.

[152, 174, 175] introduced Semiparametric Support Vector Machines (SSVM). The SSVM consists of a parametric model and a nonparametric model. The parametric model encodes prior knowledge that constrains the nonparametric model. Nguyen and Tay [121] explored the use of multiple kernels in the nonparametric model and spline functions in the parametric model to model

functions that have different properties in different portions of the function domain. We expect to include other sources of prior information. Adding prior information to Kernel Machines is an active area of research. [92, 181] provide an survey of the field, relating mostly to Support Vector classification. [143] looked at the problem of using prior knowledge to construct kernels to deal with local invariance and locality in images.

Standard decomposition methods such as [131] and SVM^{light}[80] can not deal with the multiple equality constraints that appear in SSVM. General quadratic solvers can deal with multiple constraints but don't take advantage of the structure of Support Vector Machines. As a result they are limited in the size of the problem that they can solve.

Equality constraints work well for specifying constraints at points. [106, 107] used a fundamental theorem of the alternative for convex functions to convert linear or non-linear prior knowledge into inequality constraints.

Finally, transform invariants allow incorporation of prior knowledge locally. [181] surveyed methods for incorporating prior knowledge in SVMs including local transform invariants.

4.7 Semi-Supervised Methods

[209] provides an extensive literature survey on Semi-supervised Learning. The basic idea is to exploit both the structure and a small amount of labeled data to find a good classifier. Semiparametric Support Vector Machines (SPSVM) are described in [152]. Our algorithm exploits the structure in the data using a semiparametric model.

A formulation that is similar to the nonparametric portion of our model is [144]. They used an unsupervised method to construct surfaces from point clouds of data.

4.8 Semiparametric Models

Semiparametric models account for the data at multiple scales simultaneously. Parametric models can be used to infer large-scale features of a data set but may occasionally be too smooth to account for small scale features. Non-parametric models on the other hand can model data locally very well but do poorly at modeling large scale features. Semiparametric models are a synthesis of these properties. A parametric model uses a fixed or finite number of parameters to model the data [58, 185]. The number of parameters in a nonparametric model may be either infinite [27] or in some sense proportional to the size of the problem or the number of data points [184]. Often semiparametric models consider the parametric portion of the model informative and the nonparametric portion to be a "nuisance" factor [8]. In our model both the parametric and nonparametric portions of the model are useful and important.

Our semiparametric model combines a family of implicit algebraic functions with Support Vector Machines [99, 121, 152] and extends them into the semiparametric realm.

4.9 Discussion

This chapter introduced the fundamental concepts and constructs that underlie our Semiparametric Geometric Model approach. The most important abstract elements are the definition of manifolds, foliations, implicit algebraic surfaces as foliations of a manifold, and total ordering of foliations. The most important algorithmic notions reviewed are semiparametric models, manifold learning, and semi-supervised learning.

In Chapter 5 these elements are combined and expanded to give the formulation of our new Semiparametric Geometric Model.

Chapter 5

Semiparametric Geometric Modeling

In this chapter a new modeling framework is presented called Semiparametric Geometric Modeling (SGM). An SGM brings together elements of Differential Geometry, Differential Topology and Machine Learning to model the data at multiple scales on multiple volumetric structures. The resulting algorithm enables automatically extracting and modeling white matter volumetric structures from DTI data. Next, in Chapter 6, methods for fitting the model to data are presented.

5.1 Introduction

SGM is a new model and algorithm for extracting and modeling portions of white matter volumetric structures. The result is a set of functions that describe the local internal and global geometric properties of those volumetric structures. SGM combines a Semi-supervised Manifold Learning Algorithm (SMLA) with a Semiparametric Functional Approximation Model (SFAM). We begin with the assumption that different white matter volumetric structures in the brain are described by separate manifolds each of which is a smooth representation of the data that lie on that manifold. Since the determination of which manifold the data

lie on is a function of the organization of the data the first thing to do is to take a closer look at the data.

Recall the Diffusion Tensor (DT) is made up of three orthogonal eigenvectors, $\varepsilon_{\parallel}$, $\varepsilon_{\perp}$ and $\varepsilon_{\parallel\vdash}$. $\varepsilon_{\parallel}$ points along the white matter fibers so its link to large scale geometry is clear. $\varepsilon_{\perp}$ is normal to the sheet-like organization of white matter in a given region. $\varepsilon_{\parallel\vdash}$ rounds out the description. See Figure 3.1. The eigenvectors of the DTs can be decomposed into three orthogonal vector fields. One vector field is composed of all of the $\varepsilon_{\parallel}$, the second vector field is composed of all of the $\varepsilon_{\perp}$ and the third vector field is composed of the $\varepsilon_{\parallel\vdash}$ eigenvectors. An SMLA is used to cluster tensors with similar organization and location and that fit closely to an SFAM model. An SFAM simultaneously "fits" the gradients of a global function to the Diffusion Tensor (DT) vectors. Since the SFAM only fits gradients to the tensors finding a way to localize the model to the "middle" of the data and give a robust estimate of the SFAM is important. A geometric entity called the Geometric Median Surface (GMS) is a robust estimator of the position of a parametric surface in the "middle" or deepest position of a vector field. The GMS provides the property to the SGM model.

In the next section we describe what properties of the manifolds we rely on and show how those properties fit the physical problem we are addressing. We begin by looking at this as a multi-manifold learning problem. In order to find each manifold and partition the corresponding data we further formulate this as a semi-supervised learning problem. In this case a few example voxels are designated for each of the manifolds of interest and then the SGM is computed and used to add new data points to the computed models. This process is repeated until all of voxels are assigned to manifolds or have been determined to be unassignable.

We use techniques developed in machine learning and differential geometry to construct the nonparametric portion of our model. Our parametric model leverages properties of differential topology and our new Geometric Median Surface. We begin by analyzing the geometric properties of the model.

5.2 Geometric and Algebraic Models

Small bundles of white matter fibers have an obvious geometric property. That is, they may be described locally as smooth, continuous curves that have a meaningful distance or metric associated with them. At the same time they also have a clear interpretation in terms of algebraic functions as they can be modeled as manifold curves. Such manifold curves can be described as the intersection of two algebraic implicit surfaces.

Algebraic functions have a fixed number of parameters and are considered to be parametric models. Parametric models have two important properties. First, by accounting for large-scale structure they reduce the complexity and therefore the dimensionality of a model. Second, the form of the parametric model allows us to include prior knowledge about large-scale features in the model in a straightforward manner. Unfortunately, they can be poor at handling local complexity.

Conversely, nonparametric models handle local complexity well but other than some smoothness constraints generally don't model global properties well. Prior knowledge about local features of the data can be added to nonparametric models.

A semiparametric model can be stated generically as a blend of parametric and nonparametric models. Note first that either model is a weighted sum of some number of basis functions. First, a parametric function $\mathcal{F}(\vec{x}, \vec{\theta})$ over the domain of the data points $\vec{x}$ has a fixed number of basis functions with unknown weights or parameters $\vec{\theta}$. Conversely a generic nonparametric function assigns a basis function Φ to each data point and has the same number of weights as the number of data points. This can be written compactly as an inner product of weights and basis functions, $\langle \mathbf{w}, \Phi(\vec{x}) \rangle$.

Using this notation, our semiparametric model is defined as

$$\mathbf{G}(\langle \mathbf{w}, \Phi(\vec{x}) \rangle, \mathcal{F}(\vec{x}, \vec{\theta})). \tag{5.1}$$

For brevity, we fold the parameters, $\vec{\theta}$, into the notation of the parametric model and write $\mathcal{F}(\vec{x},\vec{\theta})$ as $\mathcal{F}(\vec{x})$.

Then, for each manifold we are modeling, $\mathfrak{M}_i$,

$$\mathbf{G}(\langle \mathbf{w}, \Phi(\vec{x})\rangle, \mathcal{F}(\vec{x})) = \langle \mathbf{w}, \Phi(\vec{x})\rangle + \mathcal{F}(\vec{x}) \tag{5.2}$$

$$\mathfrak{M}_i = \langle \mathbf{w}, \Phi(\vec{x})\rangle_i + \mathcal{F}_i(x). \tag{5.3}$$

The nonparametric function $\langle \mathbf{w}, \Phi(x)\rangle$ specifies the local structure of the manifold – the differential geometry, while the parametric function $\mathcal{F}(\vec{x})$ models the large scale or global structure of the manifold – the differential topology.

Now that have we stated the general form of the model we are ready to analyze the parametric and nonparametric parts separately before recombining them into a full statement of the model at the end of the chapter.

5.2.1 Parametric Models: Manifolds and Curvilinear Coordinate Systems

Formally, let the set of white matter volumetric structures be given by

$$\mathbb{W} \equiv \bigcup \mathbb{W}_i,\ i = 1, ..., \mathbb{K}, \tag{5.4}$$

where $\mathbb{W}$ is a set of $\mathbb{K}$ white matter volumetric structures. Let the set of nonlinear manifolds be given by

$$\mathfrak{M} \equiv \bigcup \mathfrak{M}_i,\ i = 1, ..., k, \tag{5.5}$$

where $\mathfrak{M}$ is the set of nonlinear manifolds.

That is, each structure, $\mathbb{W}_i$, lies on the manifold $\mathfrak{M}_i$. We are given a set of n points $(X_1, ..., X_n \in X)$ sampled from k intersecting nonlinear manifolds and a set of k indexes (i) where $i \in 1, 2, ..., k$ is an index associated with each $\mathfrak{M}_i$. We find a set of functions $f_i(X)$ that assigns each point to one or more manifolds (in case of intersecting or overlapping structures). Specifically, each structure will be

represented by a function that maps data in the input space $\mathbb{R}^3 x S^3$ to a nonlinear manifold in $\mathbb{R}^3$.

Recalling Section 4.3, the foliation of manifolds are defined as,

$$\bigcup_{\alpha \in B} \mathfrak{F}_\alpha = \mathfrak{M} \tag{5.6}$$

$$\bigcup_{\alpha \in B} \mathcal{S}_\alpha = \mathfrak{M}^3 \tag{5.7}$$

where $\mathcal{S}_\alpha$ are the implicit surfaces that foliate the manifold and therefore partition the corresponding white matter tract, and B is the set of foliations.

Based on observed global structure [95, 186] and the definition of the Diffusion Tensor, the tensor field is treated as three mutually orthogonal vector fields:

$$\bigcup_{\alpha \in B} (\mathcal{S}_{\vec{\varepsilon}_\parallel,\alpha} \cap \mathcal{S}_{\vec{\varepsilon}_\perp,\beta} \cap \mathcal{S}_{\vec{\varepsilon}_{\Vdash},\gamma}) = \mathfrak{M}^3 \tag{5.8}$$

where $\vec{\varepsilon}_\parallel$, $\vec{\varepsilon}_\perp$, and $\vec{\varepsilon}_{\Vdash}$ are the eigenvectors of the tensor in the parallel, normal, and binormal diffusion directions respectively as defined in Section 3.1. α, β and γ are the set of foliations of the manifold in each vector field corresponding to $\vec{\varepsilon}_\parallel$, $\vec{\varepsilon}_\perp$ and $\vec{\varepsilon}_{\Vdash}$.

Given that the surfaces $\mathcal{S}$ are codimension 1, the foliation represented by Eq. 5.8 is clearly a codimension 0 foliation and therefore localizes every point on the manifold.

Returning to the global properties of the manifold we note the following theorem:

Theorem 5.1. *Dupin's theorem [153] states that if we have three families of implicit surfaces such that the surfaces of all the families foliate a manifold $\mathfrak{M}^3$, and implicit surfaces from different families intersect mutually orthogonally, then the intersection curves of surfaces from different families are lines of curvature. Additionally, at each point of intersection the tangent to the curve of intersection is a principal direction that lies in both surfaces. These surfaces naturally form*

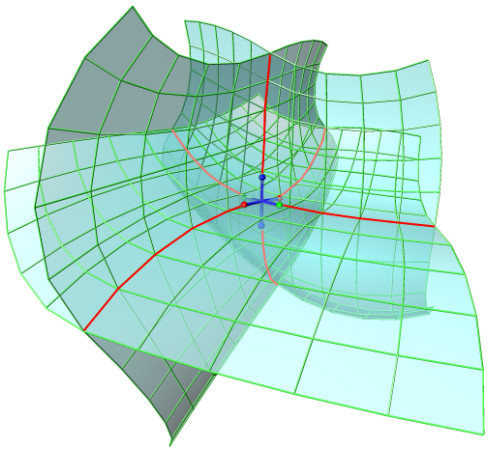

Figure 5.1: Dupin's Theorem. Blue represents the orthogonal intersection. Red are the principal lines.

a curvilinear coordinate system on $\mathfrak{M}^3$. *The curves of intersection can also be shown to be geodesics on the manifold.*

This gives us some useful properties:

1. Any point on the manifold can be uniquely defined as the intersection of three mutually orthogonal implicit surfaces (see Figure 5.2).

2. Choice of a specific point automatically gives three principal curves that are geodesics on the manifold (see Figure 5.1).

3. Given the properties of the manifold and codimension-1 foliations, it is possible to uniquely define volumes on the manifold (see Figure 5.2).

Property 2 means that for each point on the manifold the principal curve in the direction parallel to the white matter bundles is completely defined by $S_{\perp,\alpha} \cap S_{\shortparallel\vdash,\alpha}$ where α is the index of the implicit surfaces. Property 3 is the property that allows us to talk about connectivity between ROIs in terms of well-defined volumes. The function

$$\mathcal{F}(\vec{x}) = S_{\shortparallel}(\vec{x}) \cap S_{\perp}(\vec{x}) \cap S_{\shortparallel\vdash}(\vec{x}) \tag{5.9}$$

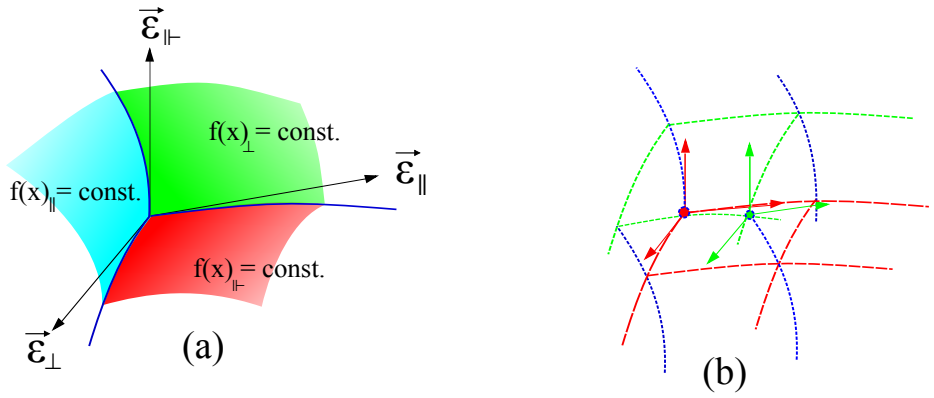

Figure 5.2: (a) Each point on the manifold is located at the intersection of three implicit surfaces. Each surface is normal to the vector field defined by ε_i. (b) As in a regular rectilinear coordinate system a volume can be defined by two points located at non-coplanar points of the "cube."

parameterizes the manifold. $\vec{x}$ are the data points.

The families of curves, or codimension 2 foliations, may now be defined as

$$\begin{aligned} \mathcal{C}_{\parallel} &= S(\vec{x})_{\perp} \cap S(\vec{x})_{\Vdash} \\ \mathcal{C}_{\perp} &= S(\vec{x})_{\parallel} \cap S(\vec{x})_{\Vdash} \\ \mathcal{C}_{\Vdash} &= S(\vec{x})_{\parallel} \cap S(\vec{x})_{\perp}. \end{aligned} \tag{5.10}$$

$\mathcal{C}_{\parallel}$ is the curve whose tangents are the $\vec{\varepsilon}_{\parallel}$ vectors. A natural question to ask is: what is the relationship between the complexity of the surfaces S_i and the complexity of the curves $\mathcal{C}_i$? Assuming two surfaces are defined as algebraic implicit functions of degree m and n by Bézout's Theorem, the degree of a curve of intersection is on the order of the product of the degrees of the surfaces or polynomials, for example degree $m \times n$. So, for example, two surfaces defined by 4^{th} degree polynomials could have curves of intersection up to order 16.

The family of curves $\mathcal{C}_{\parallel}$ have a particulary important role in the global geometry

of white matter structures. These curves are parallel to the fiber orientation and trace along the white matter fibers. Tracing these curves gives the primary method for extracting data for statistical analysis along the length of the white matter fibers.

5.2.2 Implicit Surfaces that Foliate the Manifold

An implicit function that models a surface that foliates a manifold with an associated vector field has a specific location and a specific relation to the manifold. The gradient of the implicit function is aligned point by point with the vector field throughout the volume of the manifold. Geodesics of the manifold follow paths whose tangents are also locally parallel to the vectors of the vector field. First define points on the manifold in terms of a function of the intersection of implicit functions of surfaces:

$$\mathcal{F}(\vec{x}) = \mathcal{S}(x)_{\parallel} \cap \mathcal{S}(x)_{\perp} \cap \mathcal{S}(x)_{\Vdash} \tag{5.11}$$

with the surfaces defined in terms of alignment and pairwise location on the manifold:

$$\mathcal{S}_i = 1 - \frac{\nabla f(x)_i}{\|\nabla f(x)_i\|} \cdot \vec{\varepsilon}_i + d_{\mathbb{G}}(\mathfrak{F}(\vec{\mu}), \vec{x})_{\vec{\varepsilon}_i} \tag{5.12}$$

where $\frac{\nabla f(x)_i}{\|\nabla f(x)_i\|} \cdot \vec{\varepsilon}_i$ is the alignment of the gradient of $\mathcal{S}(x)_i$ with the Diffusion Tensor eigenvectors, $\vec{\varepsilon}_i$, where $i \in (\parallel, \perp, \Vdash)$. $d_{\mathbb{G}}(\mathfrak{F}(\vec{\mu}), \vec{x})_{\vec{\varepsilon}_i}$ is a distance function related to the "location" of $\mathcal{S}_i$ on the manifold.

Distance relationships depend on the underling manifold structure. In $\mathbb{R}^3$ $d_{\mathbb{E}}(\mathfrak{F}(\vec{\mu}_i), \vec{x}_i)$ is the usual Euclidean Distance. $d_{\mathbb{K}}(\mathfrak{F}(\vec{\mu}_i), \vec{x}_i)$ is the Gaussian Kernel projection used in Principal Curves and Surfaces. Finally, $d_{\mathbb{G}}(\mathfrak{F}(\vec{\mu}_i), \vec{x}_i)$ represents the geodesic distance function from a point to the surface $\mathfrak{F}(\vec{\mu}_i)$. Distances using this measure require tracing along geodesics. Looking back to $\mathbb{R}^3$ the geodesics are straight lines and $\mathfrak{F}(\vec{\mu}_i)$ are planes so in $\mathbb{R}^3$ under those conditions $d_{\mathbb{G}}(\mathfrak{F}(\vec{\mu}_i), \vec{x}_i) = d_{\mathbb{E}}(\mathfrak{F}(\vec{\mu}_i), \vec{x}_i)$. The Gaussian Kernel projection distance function

fits data to nonparametric curves and surfaces by linear projections. Our manifolds are in general nonlinear and potentially complex so $d_{\mathbb{G}}(\mathfrak{F}(\vec{\mu}_i), \vec{x}_i)$ is determined from the data as described in Section 5.2.3.

Ultimately an estimation of the functions that describe the manifold will need to be as robust as possible in the presence of noisy data. That issue is addressed in the next section.

5.2.3 Geometric Median Surface (GMS)

There are an infinite number of implicit functions whose gradients would fit the vector data. Some means is required to localize the estimate "near" or in the "middle" of the data in order to properly constrain the function. One way to approach this kind of constraint is to look for the foliation that most evenly divides the data in a geometric sense. The function that locates this foliation is known as a depth function. [7] surveys an number of methods for robust estimation of data depth that are based on various forms of medians and subspace division techniques. Since we have requirements that those methods fail to meet we present a new depth function called the Geometric Median Surface.

Definition 5.2. ***Geometric Median Surface***
Geometric Median Surface *is the implicit surface that minimizes the sum of distances along geodesics of the manifold:*

$$S_i^M = \arg \min_{\substack{\vec{\mu}_{i,j} \in \mathfrak{F}_i(\vec{\mu},\vec{x})=0 \\ \vec{x}_j \in \mathcal{M}}} \sum_j^N d_{\mathbb{G}}(\mathfrak{F}_i(\vec{\mu_{i,j}}), \vec{x}_{i,j}). \tag{5.13}$$

More generally, the Weighted Geometric Median Surface is defined as

$$S_i^M = \arg \min_{\substack{\vec{\mu}_{i,j} \in \mathfrak{F}_i(\vec{\mu},\vec{x})=0 \\ \vec{x}_{i,j} \in \mathcal{M}}} \sum_j^N w_{i,j} d_{\mathbb{G}}(\mathfrak{F}_i(\vec{\mu}_{i,j}), \vec{x}_{i,j}) \tag{5.14}$$

where $\mathfrak{i} \in \{\|, \perp, \Vdash\}$, $\vec{\mu}$ are points that lie on the GMS and the distance function is $\mathfrak{d}_{\mathbb{G}}(\mathfrak{F}(\vec{\mu}_{\mathfrak{i}}), \vec{x}_{\mathfrak{i}})$.

The GMS is a constraint of the formulation of the implicit function that foliates the manifold of the model. As such it is also the primary object that represents the global geometry of the SGM. Given the definition of the GMS, at the intersection of the three GMSs are the three principal curves that are "deepest" in the data, $\mathcal{C}^{\mathcal{M}}_{\|}$, $\mathcal{C}^{\mathcal{M}}_{\perp}$, and $\mathcal{C}^{\mathcal{M}}_{\Vdash}$. In fact $\mathcal{C}^{\mathcal{M}}_{\|}$ plays a role similar to the medial axis. Unlike the medial axis, however, $\mathcal{C}^{\mathcal{M}}_{\|}$ is defined by the global geometry of the model that includes all of the internal structure. The medial axis is defined only relative to the surface of a structure.

5.3 One Class Semiparametric Support Vector Model (SSVM)

The SSVM is at the heart of the Multimanifold Learning algorithm. It combines it's own local similarity kernels with the semiparametric models of the SGM to decide which points to add, keep, or remove from the model.

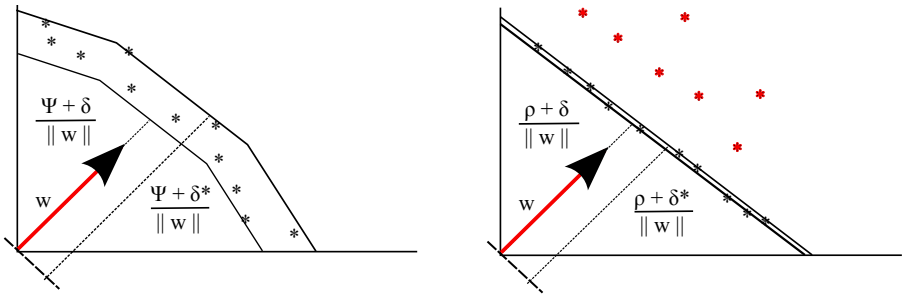

Figure 5.3: To build the model we use $\Psi = \mathfrak{d}_{\mathbb{G}}(\mathfrak{F}(\vec{\mu}, \vec{x}))$ and to evaluate the model we set $\rho = 1$, which gives us a model of the convex hull of the training data. Any new points that evaluate to the interior of the convex hull are added to the model.

Consider the following loss function:

$$V(f(x)) \equiv \begin{cases} 0 & \text{, if } \left|\frac{\tau}{2} + f(x)\right| < \frac{\tau}{2} \\ \left|\frac{\tau}{2} + f(x)\right| - \frac{\tau}{2} & \text{, otherwise.} \end{cases} \tag{5.15}$$

This leads to the following optimization problem:

$$\min_{\mathbf{w}\in\mathfrak{H},\xi^{(*)}\in\mathbb{R}^m} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{m\nu}\sum_{i=1}^{m}(\xi_i^* + \xi_i) \tag{5.16}$$

$$\text{subject to} \begin{cases} -\xi_i \leqslant \langle \mathbf{w}, \Phi(x_i)\rangle + \lambda\mathcal{F}(x_i) \leqslant \tau + \xi_i^* \\ \xi \geqslant 0 \\ \xi^* \geqslant 0 \end{cases}$$

where $\|\mathbf{w}\|^2$ is a regularization term, $\Phi(x_i)$ maps x_i into high-dimensional feature space, ξ_i^* and ξ_i are slack variables that allow some flexibility in fitting points between the hyperplanes, and $\tau = \delta^* - \delta$ is the distance between the hyperplanes (we set $\delta = 0$ and $\delta^* = \tau$). In practice, the map $\Phi(x_i)$ is unknown. The *kernel trick* allows $\Phi(x_i)$ to be used implicitly rather than explicitly for finding this mapping, i.e., $K(x_i, x_j) = \Phi^T(x_i)\Phi(x_j)$.

The Support Vector portion of our model (Section 4.7) solves a nonlinear PDE by projecting the Diffusion Tensors into a high-dimensional space. This alone doesn't guarantee a reasonable solution [157]. However, as we shall see, our method solves families of PDEs for the nonlinear case by iteratively solving what amounts to flow problems guided by the global parametric portion of the model. Each iteration includes a step that can be interpreted as applying boundary conditions to these flow problems in order to choose the next step in the look-ahead prediction of the solution trajectories of the PDEs.

5.4 Discussion

This chapter presented a new model that can automatically extract and model white matter volumetric structures from DTI data. The new model and the algorithm to compute it are called Semiparametric Geometric Modeling (SGM). SGM builds on elements of Differential Geometry, Differential Topology and Machine Learning to model the data at multiple scales on multiple volumetric structures. The next chapter formulates the problem of fitting the Semiparametric Geometric Model to data as a multi-objective optimization problem.

Chapter 6

Fitting a Semiparametric Geometric Model by Solving a Multi-Objective Optimization Problem

6.1 Introduction

In this chapter we look at how to go about fitting the SGM to data. Finding the "best" fit to the data of each part of the model and then combining the parts would very likely give a model that is a poor fit to the data. Since there are multiple parts or objectives in the model, a better approach is to combine the objectives together to form a multi-objective function. If solving each objective separately and combining them gives a good solution the solution is trivial. It is far more likely that the objectives of the model interact with each other in potentially complex ways. Fitting data to a model where the various objectives conflict or interact is known as multi-objective optimization. The multi-objective optimization problem for fitting the SGM weights the various parts of the model allowing them to be adjusted or traded off relative to one another. As a result

a whole sets of "best" solutions are described as the weights are varied. These optimal solutions are said to be Pareto Optimal [114]. It is up to the modeler to determine when a given solution is acceptable. Given this caveat fitting the SGM is formulated as the solution to a multi-objective optimization problem.

6.2 The Multi-Objective Optimization Problem

The formulation of the multi-objective function (MOF) used to fit the SGM is referred to as a scalarized MOF. This means that the optimization problem is of the form

$$\min \sum_{i=1}^{k} \lambda_i f_i(x) \tag{6.1}$$

where i indexes the objective functions $f_i(x)$ and $\lambda_i > 0$ are the weights applied to the objective functions. It's important to understand that the λ_is are inputs to the problem. The λ_is are determined by the modeler, or some wrapper function, that evaluates the desirability of a resulting solution and iteratively changes them according to outside criteria until those criteria are met.

Beginning with the semiparametric statement of the model,

$$\mathbf{G}(\vec{x}, \mathfrak{F}(\vec{x})) = \sum_{i}^{N} \alpha_i \mathbb{K}(\vec{x}, \vec{x}_i) - \mathfrak{F}(\vec{x}) \tag{6.2}$$

the scalarized MOF is:

$$\min_{\mathbf{w} \in \mathfrak{H}, \xi^{(*)} \in \mathbb{R}^m} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{m\nu} \sum_{i=1}^{m} (\xi_i^* + \xi_i) \tag{6.3}$$

$$\text{subject to} \begin{cases} -\xi_i \leqslant \langle \mathbf{w}, \Phi(x_i) \rangle + \lambda_0 \mathcal{F}(x_i) \leqslant \tau + \xi_i^* \\ \xi^{(*)} \geqslant 0, \end{cases}$$

and the parametric function is expanded out as

$$\begin{aligned}\mathfrak{F}(\vec{x}) = \min_{x\in\mathfrak{M}} \Bigg\{ & \lambda_1 \|\vec{U}_{\|} \cdot \vec{\varepsilon}_{\|}\|_{L_\infty} + \lambda_2 \left[\arg\min_{x\in\mathfrak{M}} \| d_{\mathbb{G}}(\mathcal{F}(\vec{\mu}), \vec{x})_{\vec{\varepsilon}_{\|}} \|_{L_1}\right] \\ & + \lambda_3 \|\vec{U}_{\perp} \cdot \vec{\varepsilon}_{\perp}\|_{L_\infty} + \lambda_4 \left[\arg\min_{x\in\mathfrak{M}} \| d_{\mathbb{G}}(\mathcal{F}(\vec{\mu}, \vec{x})_{\vec{\varepsilon}_{\perp}} \|_{L_1}\right] \\ & + \lambda_5 \|\vec{U}_{\shortmid\!\vdash} \cdot \vec{\varepsilon}_{\shortmid\!\vdash}\|_{L_\infty} + \lambda_6 \left[\arg\min_{x\in\mathfrak{M}} \| d_{\mathbb{G}}(\mathcal{F}(\vec{\mu}), \vec{x})_{\vec{\varepsilon}_{\shortmid\!\vdash}} \|_{L1}\right] \\ & + \lambda_7 \left\| 1 - (\vec{U}_{\|} \cdot (\vec{U}_{\perp} \times \vec{U}_{\shortmid\!\vdash})) \right\|_{L_2} \\ & + \lambda_8 \| \sum_{i,j} H(\mathcal{F}(\vec{x}))_{i,j} \|_{L_2} \Bigg\}\end{aligned} \tag{6.4}$$

where $\vec{U}_{(\cdot)}$ is the unit vector $\frac{\nabla f(x)_{(\cdot)}}{\|\nabla f(x)_{(\cdot)}\|}$ and $H(\mathcal{F})$ is the Hessian matrix with rows i and columns j.

This is a fairly large set of trade-off parameters. Fortunately, as we will see in Chapter 8, once a reasonable set of values is determined empirically, the optimization is relatively insensitive to changes in most of the weights and all but two of them were held constant for all of the models tested. A useful approach to solving this kind of multi-objective optimization problem is nonlinear least-squares. However, the norms for two object functions of the model are not compatible with the L_2 norm required for nonlinear least-squares. As a result we choose and iterative method of solution. The global MOF is solved by iteratively solving a nonlinear least-squares for the weighted objectives, while the L_1 L_∞ objective functions are solved using an iteratively reweighed least-squares approach inside the global optimization. This allows standard solvers for nonlinear least-squares to be used without modification.

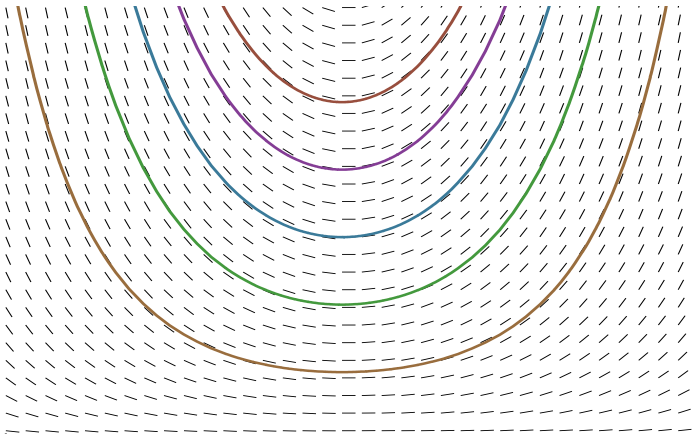

Figure 6.1: Slope or orientation field. Similar to the eigenvectors of the Diffusion Tensor, this field has orientation but not direction. The lines here are tangent to the orientations.

6.3 Solving the Multi-Objective Optimization Problem

As described in the last section the approach to solving this multi-objective optimization problem is nonlinear least-squares. The scalarized multi-objective problem is wrapped around a set of solvers for the individual objective functions. The overall algorithm is summarized next and then the solution methods for the individual objective functions are presented. See Figure 6.2 for a flowchart of the algorithm. The next section describes the algorithm for calculating the infinity norm used to align the function gradients with the tensor eigenvectors.

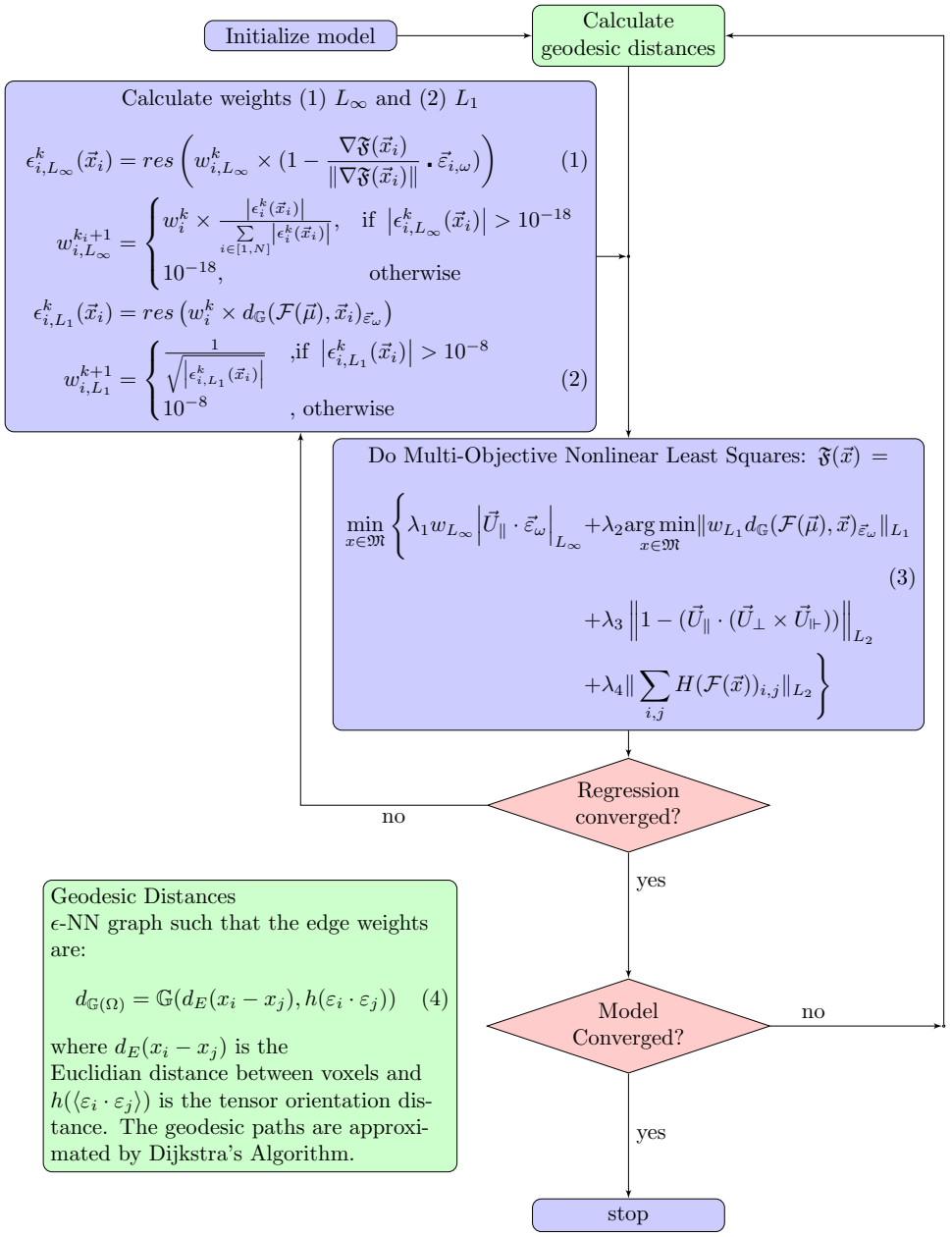

Figure 6.2: Flowchart of the algorithm for solving the multi-objective optimization problem.

6.4 Fitting L_∞ by Iteratively Reweighed Least-Squares (IRNLS)

Our model calls for matching the gradients of $\nabla\mathfrak{F}(\vec{x}_i)$ and the vector field as closely as possible. This requires minimizing

$$\min \sum_{i\in\mathfrak{M}} \left| 1 - \frac{\nabla\mathfrak{F}(\vec{x}_i)}{\|\nabla\mathfrak{F}(\vec{x}_i)\|} \cdot \vec{\varepsilon}_{i,\omega} \right|_{L_\infty}. \tag{6.5}$$

This amounts to minimizing the maximum worst fit. However, L_∞ is not equivalent to L_2 in nonlinear least-squares. So, some adjustment needs to be made to allow nonlinear least-squares to compute this norm.

One approach to making an L_∞ norm problem look like an L_2 norm so it can be solved using nonlinear least-squares is to iteratively "reweight" the residuals of the equivalent L_2 problem. The nonlinear least-squares problem is solved to completion multiple times. At each iteration the residuals from the current iteration combine weight from the previous iteration as follows:

$$\epsilon^k_{i,L_\infty}(\vec{x}_i) = res\left(w^k_{i,L_\infty} \times \left\| 1 - \frac{\nabla\mathfrak{F}(\vec{x}_i)}{\|\nabla\mathfrak{F}(\vec{x}_i)\|} \cdot \vec{\varepsilon}_{i,\omega} \right\|_{L_2} \right) \tag{6.6}$$

$$w^{k+1}_{i,L_\infty} = \begin{cases} w^k_i \times \frac{|\epsilon^k_i(\vec{x}_i)|}{\sum_{i\in[1,N]} |\epsilon^k_i(\vec{x}_i)|}, & \text{if } \left|\epsilon^k_{i,L_\infty}(\vec{x}_i)\right| > 10^{-18} \\ 10^{-18}, & \text{otherwise} \end{cases}$$

giving new weights to be used at the next iteration. Since the objective function is divided by the weights an extremely small weight can destabilize the numeric solution. Of course if a weight is very close to zero then the point that it is weighting is very close to the minimum. One approach to dealing with the stability issue would be to simply remove the point as being at the minimum. Unfortunately during the course of finding the global minimum that point could be moved away

from the minimum and it could be important bring it back into play. Rather than removing and adding the point we simply fix the smallest value that the weight can have at any iteration to 10^{-18}. This choice works well in practice. The entire loop is executed until the norm of the difference between the values of the parameters of the function change by less that 10^{-8}. At this point we haven't mentioned an important issue. Recall that the Diffusion Tensor eigenvectors have can't be interpreted to have a direction in the usual vector sense. This can't be ignored. As a result it is necessary to have a method to give the eigenvectors a direction and to be sure that the given directions result in a consistent handedness for the orthogonal eigenvectors. The next section outlines an algorithm that conditions the eigenvectors to give them these properties.

6.4.1 Conditioning the Vector Data

As described in Section 3.1, the eigenvectors of the Diffusion Tensor are orientation or slope fields, rather than true vectors, as the direction the vector is pointing is ambiguous. Finding stable functions for the Implicit Surfaces requires consistently oriented vectors, however some previous work has been done to find consistent orientations for normal vectors on surfaces defined by point clouds [76, 88]. [176] resolves the vector orientation consistency problem for surface normals by formulating it as an Ising model and finding the solution using simulated annealing. However, we are unaware of any work to date that attempts to consistently orient volumetric vector data. As a first approximation and for speed and computational simplicity, we extend the method in [76] to volumetric vector data. This involves the following:

1. Construct a Riemannian (conformal) graph of the vector field.
    a) Choose a spherical region around each point P_0 of radius γ.
    b) For all points within that sphere, calculate edge weights $w_{P_0,P_j} = 1 - \left|\langle \varepsilon_0, \varepsilon_j \rangle\right|$ where ε_j is associated with point P_j.

2. Construct a minimum spanning tree on this graph.

3. Perform a Depth-First Search on the tree.

   a) Calculate $\langle \varepsilon_i, \varepsilon_j \rangle$ for each node visited in the graph.

   b) If $\langle \varepsilon_i, \varepsilon_j \rangle < 0$ and $i < j$, "flip" the orientation of the vector of ε_j where j is the index of the newest node visited in the tree.

4. Repeat this for all three vector fields.

5. Adjust the combined vector fields to be a right-handed system.

While this is a reasonable way to construct an oriented vector field, it doesn't guarantee that all the vectors will be consistently oriented. Even more importantly, in the presence of considerable torsion, as may be found in portions of the corticospinal tract, the orientation of the vectors could be changing rapidly enough that this method will fail to produce a "good" orientation for all vectors. Fortunately, it is straightforward to detect this problem while computing the iterated least-squares solution. In other situations the orientation may genuinely be ambiguous, but we leave this for investigation at a later time.

There are an infinite number of implicit functions whose gradients would fit the vector data. Some means is required to localize the estimate "near" or in the "middle" of the data in order to properly constrain the function. One way to approach this kind of constraint is to look for the foliation that most evenly divides the data in a geometric sense. The function that locates this foliation is known as a depth function. In the next section we define a new entity that is a robust estimate of the location of the deepest surface in the data called the Weighted Geometric Median Surface.

6.5 Weighted Geometric Median Surface – Location Constraint

Definition 6.1. *The Weighted Geometric Median Surface is defined as*

$$\mathcal{S} = \arg\min_{x \in \mathfrak{M}} \sum_{i}^{N} w_i \, d_D(\mathfrak{F}(\vec{\mu}_i), \vec{x}_i) \tag{6.7}$$

where $\mathfrak{F}(\vec{\mu}_i)$ is a foliation or implicit surface that is the "deepest" surface on a manifold. The points on the surface that intersect the geodesic pathways are designated $\vec{\mu}$. d_D is a distance function based on the structure of the manifold that the data is sampled from.

For example, $d_{\mathbb{E}}(\vec{\mu}_i, \vec{x}_i)$ is Euclidean Distance. $d_{\mathbb{K}}(\vec{\mu}_i, \vec{x}_i)$ is the Gaussian Kernel projection used in Principal curves and surfaces. And $d_{\mathbb{G}}(\vec{\mu}_i, \vec{x}_i)$ is the Geodesic distance function used in our model:

$$\sum_{i \in \mathfrak{M}} \left| d_{\mathbb{G}}(\mathcal{F}(\vec{\mu}), \vec{x}_i)_{\vec{\varepsilon}_\omega} \right|_{L_1} \tag{6.8}$$

The local topology is estimated using an ϵ-NN approach. The geodesic distances are approximated by constructing a graph such that the edge weights are

$$d_{\mathbb{G}(\Omega)} = \mathbb{G}(d_E(x_i - x_j), (n_i \cdot n_j)) \tag{6.9}$$

where $d_E(x_i - x_j)$ is the Euclidian distance between voxels and $n_i \cdot n_j$ is the inner product of the normals of the data. $\mathbb{G}$ is a distance similar to [118]. The geodesic paths are then approximated by applying Dijkstra's Algorithm [44, 47].

6.5.1 Modified Weiszfeld Algorithm

A robust estimator of location in a Euclidean space is the Geometric Median or the Fermat-Weber point [7, 9, 102]. The Geometric Median doesn't take any

additional structure other than location into account. We have extended the Geometric Median to take additional structure into account resulting in what we call the **Weighted Geometric Median Surface**. This is the surface that is the "deepest" surface in a data set. Without loss of generality, we assume the weights are all 1 and discuss only the Median Geometric Surface. The Median Geometric Surface is estimated by computing a Manifold Geometric Median for each point on the surface as follows. For each point μ_i on the surface, take an ϵ region around that point and find the sum of geodesic distances to that region. Then, estimate a new point μ_i by taking a single step of the modified Weiszfeld Algorithm [188] to reduce the sum of distances. Shift μ_i by a step scaled by β:

$$\mu^{k+1} = \mu^k - \beta \frac{\left(\sum_{j=1}^{N} \frac{w_i \mu_i^k}{d_{\mathbb{G}(\mu_{i,\epsilon}^k)}} \right)}{\left(\sum_{j=1}^{N} \frac{w_i}{d_{\mathbb{G}(\mu_{i,\epsilon}^k)}} \right)}. \tag{6.10}$$

Then a new surface is estimated using μ^{k+1} in the next iteration of IRNLS using the L_1 weighting scheme described in the next section.

6.5.2 Computing L_1 using IRNLS

As with the L_∞ norm in Section 6.4 we need to modify an L_1 norm to look like an L_2 norm. Here again we use a reweighting scheme that simulates an L_1 norm.

In the following the residual for the current iteration $(k+1)$ is divided by the square root of the residual from the previous iteration (k) the solver then squares this ratio as a step to calculating its optimum value.

$$\left(\frac{\epsilon^{k+1}}{\sqrt{|\epsilon^k|}} \right)^2 = \frac{(\epsilon^{k+1})^2}{|\epsilon^k|} = |\epsilon|. \tag{6.11}$$

As a result the algorithm converges to the absolute value or the L_1 norm of the residuals instead of the L_2 norm of the residuals. At each iteration the weights are

estimated as follows:

$$\epsilon^k_{i,L_1}(\vec{x}_i) = res\left(w^k_i \times d_{\mathbb{G}}(\mathcal{F}(\vec{\mu}), \vec{x}_i)_{\vec{\varepsilon}_\omega}\right) \tag{6.12}$$

$$w^{k+1}_{i,L_1} = \begin{cases} \frac{1}{\sqrt{\left|\epsilon^k_{i,L_1}(\vec{x}_i)\right|}} & \text{, if } \left|\epsilon^k_{i,L_1}(\vec{x}_i)\right| > 10^{-8} \\ 10^{-8} & \text{, otherwise.} \end{cases} \tag{6.13}$$

Points with residuals $\epsilon \approx 0$ contribute little to the solution of the minimization. However, since we are taking the multiplicative inverse to create the new weights, we get very large numbers. Setting $\frac{1}{\epsilon}$ equal to a small number in these instances greatly improves the stability of the solution.

6.6 Orthogonality Constraint

Up to this point the formulation has dealt with elements that model each vector field independently. At this point these elements a brought together globally by requiring that the gradients of the surface functions be orthogonal at each point in the volume. When three vectors are orthogonal this means that the cross produce of any two of them gives a vector that is parallel with the third vector. Consequently, the absolute value of dot product of the third vector and and the resultant from the cross product is equal to 1. It is necessary to subtract this value from one to get a function that is 0 when the vectors are parallel. This constraint required locally by the Riemann manifold assumption and globally to fulfill the requirements of Dupin's Theorem. Recalling that $\vec{U}_{(\cdot)}$ is the unit vector $\frac{\nabla f(x)_{(\cdot)}}{\|\nabla f(x)_{(\cdot)}\|}$ with $\cdot \in \{\|, \perp, \Vdash\}$ the next term to be minimized is

$$\left\| 1 - (\vec{U}_{\|} \cdot (\vec{U}_{\perp} \times \vec{U}_{\Vdash})) \right\|_{L_2}. \tag{6.14}$$

When the residuals of this term are 0, the implicit surfaces represented by the functions are orthogonal at all points of intersection. The norm of this term is already an L_2 norm and needs no modification.

6.7 Regularization

The final term is the regularization term. This can be thought of as calculating the Hessian matrix for all points,

$$H(\mathcal{F}(\vec{x_i})_{k,l,m}) = \begin{bmatrix} \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_1^2} & \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_1 \partial x_2} & \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_1 \partial x_3} \\ \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_2 \partial x_1} & \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_2^2} & \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_2 \partial x_3} \\ \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_3 \partial x_1} & \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_3 \partial x_2} & \frac{\partial^2 \mathcal{F}(\vec{x_i})}{\partial x_3^2} \end{bmatrix}, \tag{6.15}$$

and summing the results:

$$\sum_{i \in \mathfrak{M}} \left| \sum_{k,l,m \in H} H(\mathcal{F}(\vec{x}_i))_{k,l,m} \right|_{L_2}. \tag{6.16}$$

This term is similar to the energy term used in thin plate splines [178]. The "rougher" $\mathcal{F}$ is, the larger this term becomes. Conversely, this term decreases as $\mathcal{F}$ becomes smoother. Keeping $\mathcal{F}$ smooth while still allowing it to fit the data well is the motivation for using this term. As with the orthogonality constraint, the norm of this term is already an L_2 norm and needs no modification.

6.8 Semi-Supervised Learning and the Semiparametric SVM

All elements of the parametric portion of the SGM are now in a form that can be solved as a Multi-Objective Nonlinear Least-Squares problem using a trust region reflexive method implemented in the *lsqnonlin* function in Matlab [110]:

$$\min_{\mathbf{w}\in\mathfrak{H},\xi^{(*)}\in\mathbb{R}^m} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{m\nu}\sum_{i=1}^{m}(\xi_i^* + \xi_i) \tag{6.17}$$

$$\text{subject to} \begin{cases} -\xi_i \leqslant \langle \mathbf{w}, \Phi(x_i)\rangle + \lambda\mathcal{F}(x_i) \leqslant \tau + \xi_i^* \\ \xi^{(*)} \geqslant 0 \end{cases}$$

where $\|\mathbf{w}\|^2$ is a regularization term, $\Phi(x_i)$ maps x_i into high-dimensional feature space, ξ_i^* and ξ_i are slack variables that allow some flexibility in fitting points between the hyperplanes, and $\tau = \delta^* - \delta$ is the distance between the hyperplanes (we set $\delta = 0$ and $\delta^* = \tau$). In practice, the map $\Phi(x_i)$ is unknown. The *kernel trick* allows $\Phi(x_i)$ to be used implicitly rather than explicitly when finding this mapping, i.e., $K(x_i, x_j) = \Phi^T(x_i)\Phi(x_j)$.

6.8.1 Semiparametric Slab SVM

The similarity metric is a product of five kernels. Each kernel defines a similarity with a different component (or orientation) of the Diffusion Tensor (θ, ϕ, ψ), with the distance between points ($\|x - x_i\|_{L_2}$), and invariant ($\pi \in \{FA, MD, AD, ADC, ...\}$). Each similarity kernel is a compact radial basis function and the final kernel is constructed by taking the product of the similarity kernels. The individual kernels are:

$$K_1(x, x_i) = k(-\gamma_1 * (wt_d + \|x - x_i\|_{L_2})) \tag{6.18}$$

$$K_2(x, x_i) = k(-\gamma_2 * (1 - |\langle \theta, \theta_i\rangle|)) \tag{6.19}$$

$$K_3(x, x_i) = k(-\gamma_3 * (1 - |\langle \phi, \phi_i\rangle|)) \tag{6.20}$$

$$K_4(x, x_i) = k(-\gamma_4 * (1 - |\langle \psi, \psi_i\rangle|)) \tag{6.21}$$

$$K_5(x, x_i) = k(-\gamma_5 * (1 - |\langle \pi, \pi_i\rangle|)) \tag{6.22}$$

The product of these kernels,

$$\mathbb{K}(x, x_i) = \prod_{j=1}^{5} K_j(x, x_i) \tag{6.23}$$

is the kernel used in the SVM. Since we model the local structure using kernels, we need to keep the support of the kernels local. To this end, we use radial basis functions with compact support [197]. This is similar in spirit to using k-NN or ϵ-NN to model local structure in other manifold learning algorithms [139, 163]. Note, for example the similarity to the model in [144], which was formulated as an unsupervised regression model for estimating an implicit function.

At each iteration of the optimization problem the SVM is solved using the previous iteration's data and the current iteration's estimate of $\mathcal{F}(x_i)$. Then $\mathcal{F}$ is set equal to ρ and the data are evaluated using $C(x) = \sum_{i=1}^{N} \alpha_i \mathbb{K}(\vec{x}, \vec{x}_i)$. Points for which $C(\vec{x}) > 0$ are in the affine hull of the Support Vector solution set. These points are similar to the current training data based on the kernel similarity, and are consistent with $\mathcal{F}$ as the equality constraints on the SVM. These points are then added to the points considered to be part of the extracted model. Note: a given point may be added, removed, or removed permanently as the model converges.

6.9 Discussion

This chapter described the multi-objective optimization problem and presented the algorithm for fitting an SGM to data. The scalarized multi-objective optimization problem was presented as was the concept of Pareto Optimal [114]. The individual objective functions were reviewed and algorithms for fitting the data were outlined. Finally, an algorithm for computing the full scalarized multi-objective optimization function that fits the SGM to the tensor data was presented and the importance of the modeler in determining when a given solution is acceptable was emphasized.

In the next chapter, Chapter 7, the specific elements needed to construct an

SGM for modeling white matter volumetric structures is laid out.

Chapter 7

Deriving Features and Extracting Structural Information from the SGM

7.1 Introduction: Traversing the Manifold

Many questions about white matter structure involve its geometric properties. This chapter considers the use of the SGM for exploring white matter volumetric structures geometrically and statistically. Recall that DTI data is collected using a 3D rectilinear grid. This grid bears no relationship to the global geometry of the white matter structures that it is samples. The SGM takes the DTI data as input and produces a global geometric model that reflects the global geometry of the white matter structures that generated the DTI data, and produces a curvilinear coordinate system that can be navigated in a number of ways to allow analysis that could not easily be accomplished otherwise.

This chapter is divided into five parts that walk through the properties and uses of the SGM. First, how to use the SGM to generate submanifolds that become the framework for extracting structural information from the manifold for analyzing white matter properties. Second, some examples are given that show how to

use the SGM to map points in $\mathfrak{M}^3 \mapsto \mathbb{R}^3$ and points in $\mathbb{R}^3 \mapsto \mathfrak{M}^3$. Third, we describe how to use the SGM to extract intrinsic invariant geometric properties of points on the manifold. Fourth, white matter connectivity is determined in terms of properties of the SGM that allow the global partitioning of the data into geometrically well-defined regions. Fifth, orientation statistics [109] are presented as a means for testing how well the SGM fits the data and ways are proposed for exploring small scale structure that might not be well described by the SGM. Finally, a method for performing rigid body transforms on the SGM by linear operations on the coefficients of SGM polynomials is presented. This is a first step in a registration algorithm that is outlined in the final chapter of this thesis.

7.2 Generating Submanifolds of the SGM

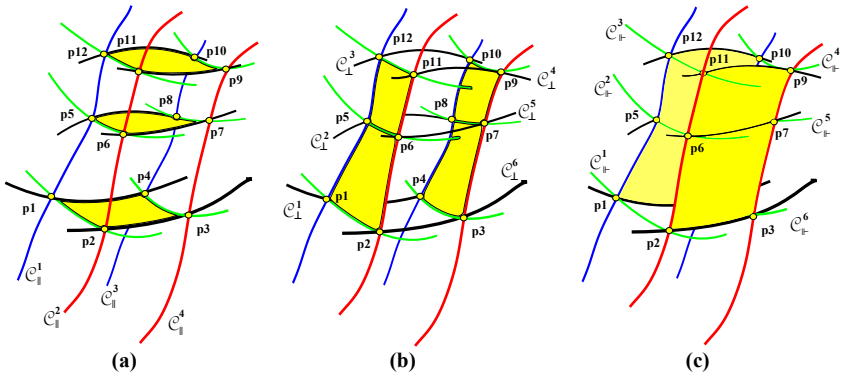

Figure 7.1: Curves generated by intersecting surfaces.

Submanifolds of the SGM consist of points, curves, and surfaces. Recall that the SGM is composed of families of orthogonal surfaces that foliate a manifold. Each surface is a submanifold of $\mathfrak{M}^3$:

$$\mathfrak{M}(\nu_{\|}, \nu_{\perp}, \nu_{\Vdash}) = S_{\|}(\vec{x}) \cap S_{\perp}(\vec{x}) \cap S_{\Vdash}(\vec{x}) \tag{7.1}$$

where $\nu_{\|}, \nu_{\perp}, \nu_{\Vdash}$ may be thought of as the intrinsic coordinates of the manifold and $\vec{x}$ are the extrinsic coordinates in $\mathbb{R}^3$ of the implicit functions. Each implicit surface $S'(\vec{x})_{\|}$ is defined at a given point $x_0 \in \mathbb{R}^3$ as

$$\nu'_{\|} = S(\vec{x_0})_{\|}, \tag{7.2}$$
$$S'_{\|}(\vec{x}) = S(\vec{x})_{\|} - \nu'_{\|} = 0,$$
$$\nu'_{\perp} = S(\vec{x_0})_{\perp}, \tag{7.3}$$
$$S'_{\perp}(\vec{x}) = S(\vec{x})_{\perp} - \nu'_{\perp} = 0,$$
$$\nu'_{\Vdash} = S(\vec{x_0})_{\Vdash}, \tag{7.4}$$
$$S'_{\Vdash}(\vec{x}) = S(\vec{x})_{\Vdash} - \nu'_{\Vdash} = 0.$$

Similar to the planes that define the Cartesian coordinate system in $\mathbb{R}^3$, the intersection of these orthogonal surfaces define a codimension 2 submanifold that in this case are points on the manifold:

$$\mathcal{P}(\nu_{\|}, \nu_{\perp}, \nu_{\Vdash}) = S'_{\|}(\vec{x}) \cap S'_{\perp}(\vec{x}) \cap S'_{\Vdash}(\vec{x}) \tag{7.5}$$

On the other hand, taking the orthogonal surfaces pairwise defines codimension 1 submanifolds that correspond to curves on the manifold:

$$\mathcal{C}'_{\|}(t) = \mathfrak{M}(\nu_{\|}(t), \nu'_{\perp}, \nu'_{\Vdash}) \tag{7.6}$$
$$= S_{\|}(\vec{x}, t) \cap S'_{\perp}(\vec{x}) \cap S'_{\Vdash}(\vec{x})$$
$$\mathcal{C}'_{\perp}(t) = \mathfrak{M}(\nu'_{\|}, \nu_{\perp}(t), \nu'_{\Vdash}) \tag{7.7}$$
$$= S'_{\|}(\vec{x}) \cap S_{\perp}(\vec{x}, t) \cap S'_{\Vdash}(\vec{x})$$
$$\mathcal{C}'_{\Vdash}(t) = \mathfrak{M}(\nu'_{\|}, \nu'_{\perp}, \nu_{\Vdash}(t)) \tag{7.8}$$
$$= S'_{\|}(\vec{x}) \cap S'_{\perp}(\vec{x}) \cap S_{\Vdash}(\vec{x}, t)$$

where as in Equations 7.2, 7.3, and 7.4, $\nu'_{\|}, \nu'_{\perp}, \nu'_{\Vdash}$ are constants or coordinates on the manifold and $\nu_{\|}, \nu_{\perp}, \nu_{\Vdash}$ are variables. t is a monotonic, non-constant

parameter of the manifold curve $\mathcal{C}$. See Figure 7.1.

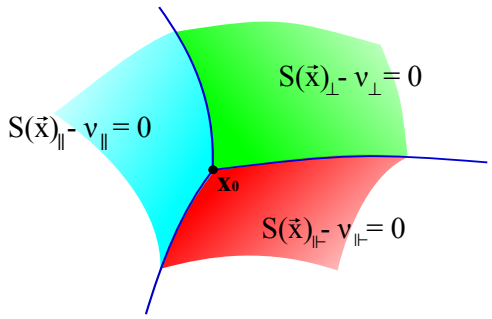

Figure 7.2: Coordinates generated by intersecting surfaces. $\nu_{\|}, \nu_{\perp}, \text{and} \nu_{\Vdash}$ are defined in Equations 7.2, 7.3 and 7.4, respectively.

7.3 Mapping from Image Space to Manifold Space

The equations in the last section are mappings from $\mathbb{R}^3$ to the manifold space $\mathfrak{M}^3$. The inverse map allows sampling the manifold $\mathfrak{M}^3$ using the diffusion data collected on a rectilinear grid in $\mathbb{R}^3$. This results in a data set that is arranged in a manner that reflects the structure of the submanifold. Scalar data derived from DT imaging such as FA, MD and RA (see Equations 3.3, 3.2, and 3.4) or tensor shape (Figure 7.9) are sampled from the 3D grid of image data. The data may also be vectorial. These invariants each give different information about tissue properties at a given voxel so it makes sense to organize them as a vector of features at each structural point for subsequent analysis.

Bringing together Equations 7.2, 7.3, and 7.4 we see that mapping the coordinates of points in $\mathbb{R}^3$ to points in $\mathfrak{M}^3$ is equivalent to finding the intersection of the level surfaces at the given point in $\mathfrak{M}^3$. This amounts to evaluating each surface function at the point x_0 giving the manifold coordinates, $\nu'_{\|}$, $\nu'_{\perp}$, $\nu'_{\Vdash}$. The inverse mapping from $\mathfrak{M}^3 \mapsto \mathbb{R}^3$ given $\nu'_{\|}, \nu'_{\perp}, \nu'_{\Vdash}$ is the solution of the system of

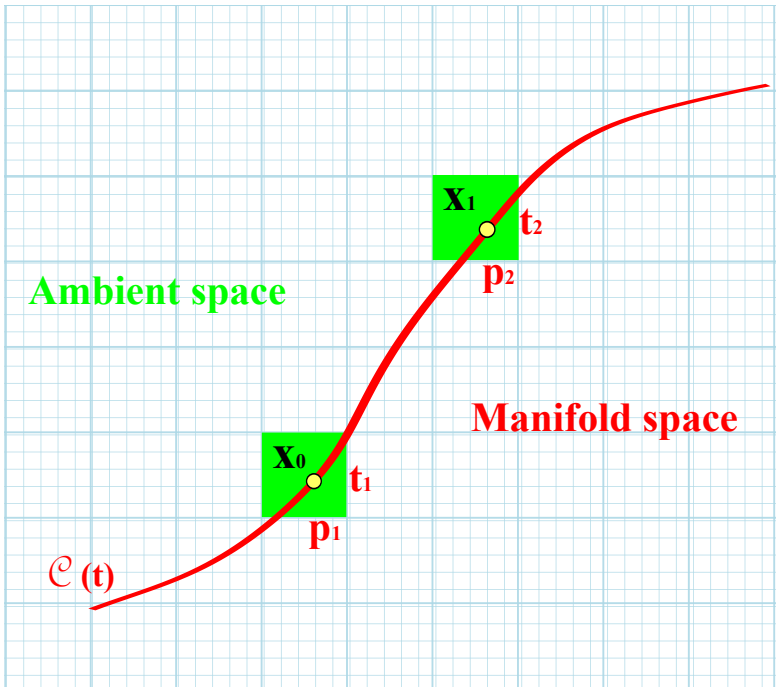

Figure 7.3: Back-sampling.
$p_1 = \mathcal{P}(\nu^1_{\parallel}, \nu^1_{\perp}, \nu^1_{\shortmid\!\vdash})$, $p_2 = \mathcal{P}(\nu^2_{\parallel}, \nu^2_{\perp}, \nu^2_{\shortmid\!\vdash})$ and $\mathcal{C}(t)$ are in the manifold space. $\vec{x}_0$ and $\vec{x}_1$ are points (ambient space) or voxels (image space). The small circles are points along $\mathcal{C}(t)$ where the voxels ($\vec{x}_0$ and $\vec{x}_1$) are mapped to the points (p_1 and p_2) in manifold space. The green squares represent scalar values that are mapped from one space to the other.

simultaneous nonlinear equations:

$$\begin{aligned} S_{\parallel}(\vec{x}_\beta) - \nu'_{\parallel} &= 0, \\ S_{\perp}(\vec{x}_\beta) - \nu'_{\perp} &= 0, \\ S_{\shortmid\!\vdash}(\vec{x}_\beta) - \nu'_{\shortmid\!\vdash} &= 0, \end{aligned} \tag{7.9}$$

for $\vec{x}_\beta$. The value of the sampled data at x_β is associated with the manifold at coordinates $\nu'_{\parallel}, \nu'_{\perp}, \nu'_{\shortmid\!\vdash}$. The procedure for mapping data in voxel space to manifold space involves traversing the submanifold (along geodesics for curves

and surfaces) of the SGM and back-sampling (Figure 7.3)

In the next three sections the machinery developed in the previous sections is used to sample the manifold in a variety of ways depending upon how we plan to analyze the data. The primary aim is to describe how to sample the manifold for use in Functional Data Analysis.

7.3.1 Sampling along Curves on the SGM

One method of sampling the manifold is to generate multiple curves on the manifold and sample them where they intersect common orthogonal surfaces. For example, Figure 7.4 depicts three curves that are used to sample the manifold where they intersect with eight orthogonal surfaces.

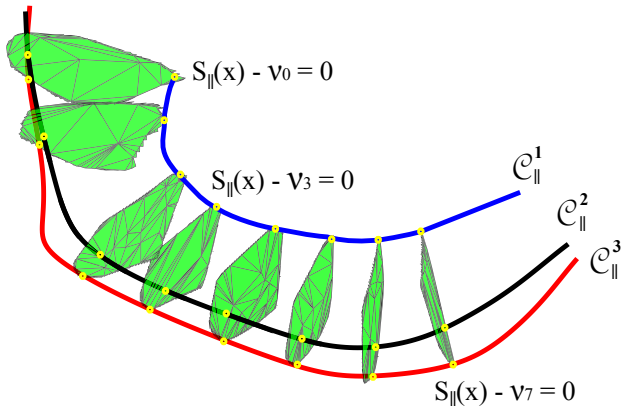

Figure 7.4: Three curves and eight surfaces produced by an SGM generated by data from the corpus callosum of a test subject.

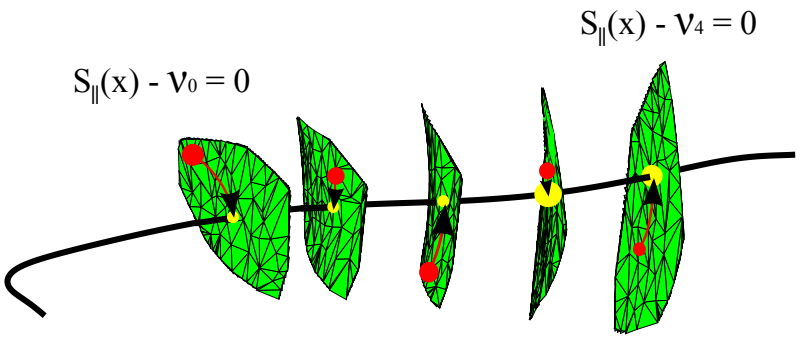

Figure 7.5: Data on orthogonal surfaces summarized on a single curve. For example, assigning the maximum FA or average FA on a surface to the curve at the point of intersection.

7.3.2 Summarizing Volumetric Structure using a Single Manifold Curve

Figure 7.5 depicts another method for sampling the manifold. This method assumes the volumetric structure is roughly tubular and that projecting a measure, for example the maximum FA value, on each surface to a single common curve is sufficient to represent the statistics of interest on the manifold.

7.3.3 Sampling by Projecting onto a Surface along Orthogonal Manifold Curves

Figure 7.7 depicts a volumetric structure that is more bilaterally extended than the previous example. In this case a surface is chosen, for example the Median Surface $S_{\perp}^{M}$ (see Equation 5.13), and measures are projected along the orthogonal curves to the surface. Point-based statistics may then be used to analyze the surfaces.

Alternatively, the $S^M_\perp$ may be sampled using curves tangent to and embedded in the surface, and then analyzed by FDA. Given the surface $S^M_\perp$, the family of curves sampling the surface are given by

$$\begin{aligned} \mathcal{C}_\parallel(t,s) &= \mathfrak{M}(\nu_\parallel(t), \nu'_\perp, \nu_{\shortparallel\vdash}(s)) \\ &= S_\parallel(\vec{x},t) \cap S'_\perp(\vec{x}) \cap S_{\shortparallel\vdash}(\vec{x},s). \end{aligned} \tag{7.10}$$

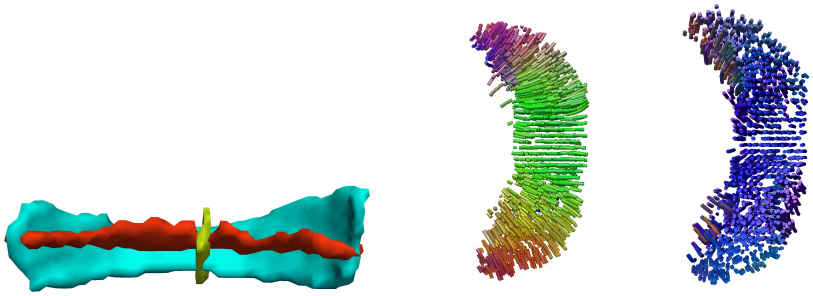

Figure 7.6: An example of the Median Surfaces derived by fitting an SGM model of the genu of the corpus callosum. The surface depicted in the left image are Green: Median Surface orthogonal to $\mathcal{C}_\parallel$, Blue: Median surface orthogonal to $\mathcal{C}_\perp$, and Red: Median orthogonal to $\mathcal{C}_{\shortparallel\vdash}$. The center image shows the curves that project data onto the blue surface, and the right image shows the curves that project data onto the red surface.

An example of three Median Surfaces, $S^M_\parallel(\vec{x})$, $S^M_\perp(\vec{x})$, and $S^M_{\shortparallel\vdash}(\vec{x})$, extracted using the SGM is shown in Figure 7.6.

In this section establishes a method for traversing the manifold and sampling the voxels in image space. The resulting data are consequently organized in a manner that reflects the organization of the structure of the manifold. In the next section the same machinery is used to traverse the manifold, but this time instead of sampling data from image space the SGM allows the sampling of position-independent geometric properties directly from the model.

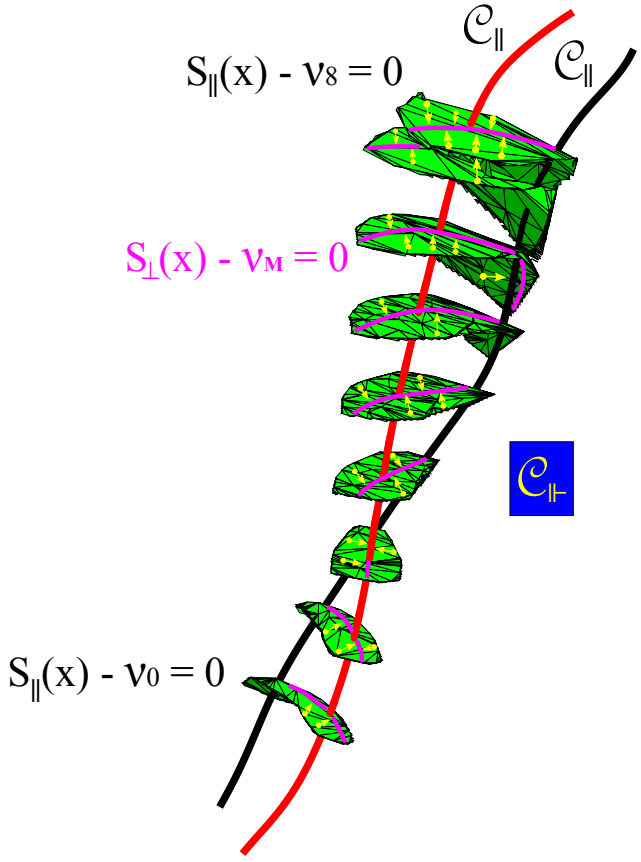

Figure 7.7: Sampling and summarizing data on the $S_{\perp}^{M}(\vec{x}) = S_{\perp}(\vec{x}) - \nu_M = 0$ median surface. The magenta lines on the $S_{\parallel}$ surfaces depict the intersection of those surfaces with $S_{\perp}^{M}(x)$ median surfaces. Tensor data are then projected along the $\mathcal{C}_{\parallel\vdash}$ (yellow) curves to the $S_{\perp}^{M}$ surface. Once the data are projected to the $S_{\perp}^{M}(\vec{x})$ surface they may be organized according to the needs of a given analysis method. Some examples of organization include a vector of data values ordered by position along the curve and the maximum or average of data values along each curve.

7.4 Geometric Properties at Points on Implicit Surfaces

We begin this section by looking at the geometric quantities that can be calculated in closed form at each point based on the polynomials in the SGM. Letting $F \in \{S_{\|}, S_{\perp}, S_{\Vdash}\}$, we start with a few preliminary definitions and then go on to give closed form formulas for the three definitions of curvature and the formula for torsion.

We begin with a description of our notation and a few definitions that allow us to write the subsequent formulas more compactly. Preliminary definitions:

$$\nabla F = \left(\frac{\partial F}{\partial x} \frac{\partial F}{\partial y} \frac{\partial F}{\partial z} \right) = (F_x F_y F_z) \tag{7.11}$$

$$H(F) = \nabla(\nabla F) \tag{7.12}$$

$$H(F) = \begin{pmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial x \partial z} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y \partial z} \\ \frac{\partial^2 F}{\partial z \partial x} & \frac{\partial^2 F}{\partial z \partial y} & \frac{\partial^2 F}{\partial z^2} \end{pmatrix}$$

$$H^*(F) = \begin{pmatrix} F_{yy}F_{zz} - F_{yz}F_{zy} & F_{yz}F_{zx} - F_{yx}F_{zz} & F_{yx}F_{zy} - F_{yy}F_{zx} \\ F_{xz}F_{zy} - F_{xy}F_{zz} & F_{xx}F_{zz} - F_{xz}F_{zx} & F_{xy}F_{zx} - F_{xx}F_{zy} \\ F_{xy}F_{yz} - F_{xz}F_{yy} & F_{yx}F_{xz} - F_{xx}F_{yz} & F_{xx}F_{yy} - F_{xy}F_{yx} \end{pmatrix} \tag{7.13}$$

Equation 7.11 gives the notation for the gradient of the function F. Equation 7.12 is the Hessian and Equation 7.13 is the adjoint of the Hessian. Next formulas for curvature involving single surfaces only are presented, followed by formulas based on manifold curves that depend on the intersection of pairs of surfaces.

7.4.1 Mean Curvature from Surfaces

Mean curvature is defined as the mean of the principal curvatures at a point on a surface and is an extrinsic measure of surface curvature:

$$K_M = \frac{\nabla F \cdot H^*(F) \cdot \nabla F^T - |\nabla F|^2 Trace(H)}{2|\nabla F|^2} \tag{7.14}$$

7.4.2 Gaussian and Principal Curvatures from Surfaces

Gaussian curvature is an intrinsic invariant whose value is not dependent on the ambient space. For each surface at a given point, the Gaussian curvature is

$$K_G = \frac{\nabla F \cdot H^*(F) \cdot \nabla F^T}{|\nabla F|^4}. \tag{7.15}$$

In order to further characterize the local geometry we need to calculate the principal curvatures. Given the mean and Gaussian curvature, the principal curvatures are

$$k_1, k_2 = K_M \pm \sqrt{K_M^2 - K_G}. \tag{7.16}$$

If k_1 and k_2 are both positive we can interpret the Gaussian to be measuring the amount of local expansion. If k_1 and k_2 are both negative we can interpret the Gaussian to be measuring the amount of local contraction. If the signs are not the same, the surface is locally hyperbolic. If $k_1 = k_2$ the point is called umbilic. This should happen only at isolated points. If one principal curvature is 0, the point is parabolic.

7.4.3 Curvature and Torsion from Implicit Curves

Curvature and torsion formulas presented in [63] can be used to calculate local differential properties of the surfaces determined by the SGM. Torsion is a particularly interesting property when analyzing the corticospinal and cortico-bulbar or cortico-pontine tracts. These regions are known to undergo considerable

torsion as the tract is traversed from the inferior to superior direction along the $C_{\parallel}$ curves of the SGM.

7.4.4 Curvature from Implicit Curves

The curvature at a given point is determined by the choice of surfaces determined by the SGM. For example, choosing $\mathcal{S}(\vec{x})_{\parallel} \cap \mathcal{S}(\vec{x})_{\perp}$, then

$$k = \frac{|((\nabla\mathcal{S}(\vec{x})_{\parallel} \times \nabla\mathcal{S}(\vec{x})_{\perp}) \cdot \nabla(\nabla\mathcal{S}(\vec{x})_{\parallel} \times \nabla\mathcal{S}(\vec{x})_{\perp})) \times (\nabla\mathcal{S}(\vec{x})_{\parallel} \times \nabla\mathcal{S}(\vec{x})_{\perp})|}{|\nabla\mathcal{S}(\vec{x})_{\parallel} \times \nabla\mathcal{S}(\vec{x})_{\perp}|^3} \tag{7.17}$$

An example of a curved volumetric structure is shown in Figure 7.8 (a).

7.4.5 Torsion from Implicit Curves

The torsion at a given point is determined by the choice of surfaces determined by the SGM. Choosing $\mathcal{S}(\vec{x})_{\perp} \cap \mathcal{S}(\vec{x})_{\shortparallel}$, then

$$\tau = \frac{Det(T_1 T_2 T_3)}{|T_1 T_2|^2} \tag{7.18}$$

where

$$\begin{aligned}
T_1 &= \nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel}, \\
T_2 &= (\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel}) \cdot \nabla(\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel}), \\
T_3 &= (\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel}) \cdot \nabla(\nabla(\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel})) \cdot (\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel})^T \\
&\quad + (\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel}) \cdot \nabla(\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel}) \cdot \nabla(\nabla\mathcal{S}(\vec{x})_{\perp} \times \nabla\mathcal{S}(\vec{x})_{\shortparallel}).
\end{aligned}$$

Figure 7.8 (b) shows an example from the experimental data of a volume that has significant torsion. As can be seen, the torsion represents the rate of change in the binormal or tertiary diffusion direction. See [51, 63, 155] for an analysis of

the derivation of these and other closed form formulas and for a more in-depth description of the geometric properties of implicit curves and surfaces.

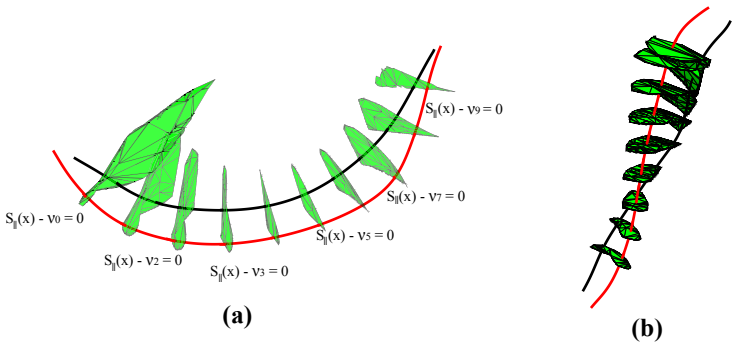

Figure 7.8: Volumes with orthogonal sampling surfaces that contain (a) high curvature with low torsion and (b) low curvature with high torsion.

7.5 Rigid Body Transformation of the SGM

We briefly mention a final property of implicit polynomial representation. Recalling Definition 4.3, **compactification** of $\mathbb{R}^3$ is accomplished by adding a point at infinity. This means transforming the problem into the real projective space $\mathbb{P}^3$. When necessary, the implicit functions representing the foliations of the manifold are reformulated using homogeneous coordinates. The resulting manifold is both compact and closed. The compactification of the SGM allows the SGM to be translated and rotated by a set of linear transforms on the coefficients of the SGM polynomials [164]. This means we can bring models describing volumetric structures from different data sets into rigid alignment. Finally, the SGM is transformed from $P\mathbb{R}^3 \mapsto \mathbb{R}^3$ by converting from homogenous coordinates back to real coordinates. Analysis and exploration of this method is part of our future work and is the first step in a proposed nonlinear registration method based on the SGM.

7.6 Estimating Connectivity

One question anatomists would like to ask is: given fibers originating in Region A, what proportion of those fibers terminate in region B? Another way to consider this question is: how can we geometrically partition the fibers into those that connect regions A and B and those begin in A and do not end in B? Probabilistic approaches can estimate a likelihood that a given voxel is on a path connecting two regions but that doesn't lead to an answer to this question. Even deterministic approaches to estimating connectivity are essentially probabilistic because streamline trajectories become more and more uncertain as the distance between the regions increases so that counting streamlines or estimating streamline densities have the same problem [14, 94].

Figure 7.11 illustrates this problem on the boundary of a connected volume of fibers. The volume is defined as the intersection of the fibers that connect the two ROIs. The cube in the figure represents a voxel whose fiber density we wish to estimate. The red-shaded region is the portion of the cube that lies inside the volume. Ideally, all of the streamlines passing through the cube and connecting the ROIs will pass through the red-shaded region. Only streamlines that originate within the fiber volume should terminate in the ROIs. For fibers passing through the cube the fiber density is then proportional to the volume of the shaded region divided by the volume of the cube. The same is true of cubes that lie entirely in the fiber volume. However, due to propagation uncertainties that, as mentioned above, increase with the distance between ROIs, these assumptions are often violated.

For example, in Figure 7.11 streamline 1 originates outside the fiber volume but passes through the external portion of the cube and terminates in one of the ROIs. Streamline 2 begins and ends outside ROIs but passes through the shaded region. Streamline 3 begins and ends in the ROIs but passes through the external portion of the cube. Streamline 4 begins in an ROI and passes through the shaded region but fails to terminate in the other ROI. Only streamline 5 begins and ends in both ROIs and remains entirely in the interior of the volume. The same analysis applies for voxels that lie entirely in the volume. An approach to dealing with

these problems is to estimate a fiber density per voxel. Many streamlines are generated in each voxel and those that terminate in both ROIs are counted. An *ad hoc* normalization is then applied to deal with the distance-related uncertainty.

Our approach leads to geometrically well-defined volumes that are fully deterministic. To define a volume using the SGM, a boundary must be defined. This is straightforward using the SGM and the curves that have been extracted, as follows. Choose any orthogonal surface $S_{\parallel} - \nu'_{\parallel} = 0$. Formulate a parametric curve, $b(t) = \{\nu_{\perp}(t), \nu_{\shortparallel}(t)\}$. This curve is determined by stepping "around" the edge of the orthogonal surface. For example Figure 7.9 shows two surfaces with blue boundary traces. Either curve may be used to estimate the boundary. In fact, since a given $C_{\parallel}(t)$ curve on the surface may be used as a common starting point, multiple surfaces may be used to give a robust estimate of the boundary. Figure 7.10 has its entire surface defined by the red curve. The total volume is then the volume of the voxels inside the surfaces defined by the boundary surface and the intersection of the volume with the ROIs. It is important to keep in mind that the surfaces are defined at the subvoxel level as is the resulting volume.

With this volume in hand we can then directly address the question posed at the beginning of this section. The volume determined as described above is divided by a volume determined by only one of the chosen ROIs. This is proportional to the number of fibers originating in, for example, ROI A that also terminate in ROI B divided by the volume of all of the fibers originating in ROI A.

Another complication needs to be considered. Nothing in the SGM requires that the volume is uniformly dense. In fact, Figure 7.12 shows three samples from SGM models of the right corticospinal tract from our experimental data. All three are densely sampled. Clearly the left example does not entirely fill the space inside the enclosed volume. Fortunately, given that the volume is well-defined, we can estimate the portion of the volume that is "empty" using Monte Carlo methods. Since all volumes may be estimated in this fashion, the ratio of the volumes is still well defined.

Considering the geometric properties, this analysis implies that the SGM is not

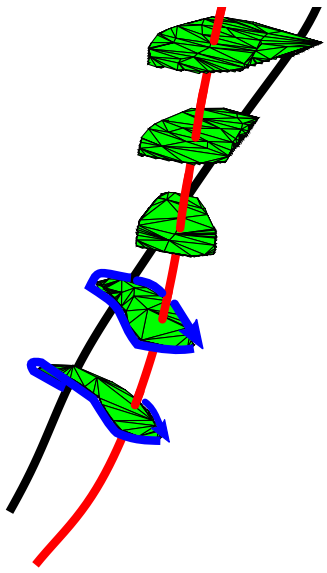

Figure 7.9: Blue lines around two of the level surfaces trace out their boundaries. Any boundary of any level surface defines the surface of the volume modeled by an SGM.

necessarily genus zero. This is an important property. For example, the presence of a tumor or even normal tissue such as deep gray matter, could result in "holes" in the manifold. Obviously this is also true for bifurcating volumetric structures. We mention this here for completeness but leave it to future work for further analysis.

There is one additional advantage of that SGM has over density or probability based connectivity. The same techniques that were described above for sampling along curves through a volume can be applied to estimating the what we call the quality of the connectivity. The quality of the connectivity is estimated by integrating the FA, MD or other scalar values along the curves through a volume. We leave the precise definition to later work.

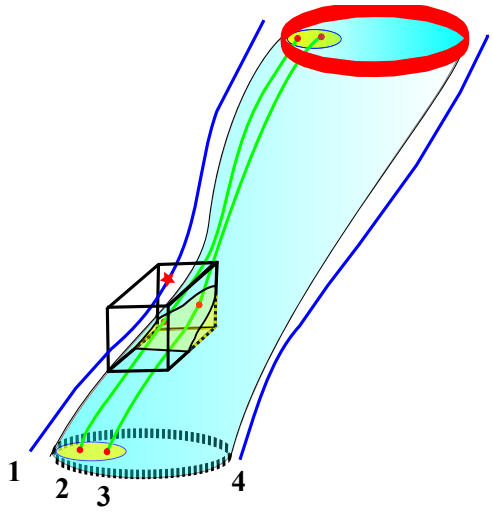

Figure 7.10: This figure illustrates the properties of an SGM volume. Curves 1 and 4 begin and end outside of the volume and can never enter the volume. Curves 2 and 3 start and end inside the volume and can never exit the volume. The test box shown in black intersects the volume in the red-shaded region to help distinguish between the interior and exterior of the volume.

7.7 Orientation Statistics

One of the questions we would like answered about the SGM is how well it fits the data. Is the SGM too smooth to properly represent the underlying volumetric structure? If the fit is good overall, are there areas where the model doesn't fit well and something needs to be done locally? The only fit we can really measure is how well the local gradients match the tensor orientations. This comparison falls under the aegis of directional statistics. If the SGM fits the tensor orientations well we would expect the gradients to fall near the mean of the local orientation distribution without the variance being too high. For example, the SGM may fit the tensor orientations in a region well but there are subregions with significant variability or bias. This may indicate either local changes in structure that may be useful in refining the model or may reflect local pathology of the tissues. For a general outline of the principles of orientation statistics see [53, 109]. We leave

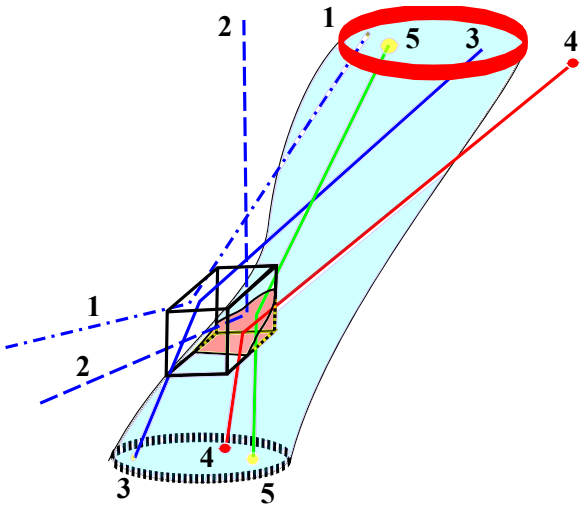

Figure 7.11: A bundle of streamlines associated with a volume. Examples of different streamlines that enter and leave the volume are shown. Due to noise and other factors it is difficult to determine how well the two ROIs are connected by the streamlines. The usual practice is to generate many streamlines, make adjustments for how far the ROIs are from each other, and define the connection as a density estimate. The test box shown in black intersects the volume in the red-shaded region to help distinguish between the interior and exterior of the volume.

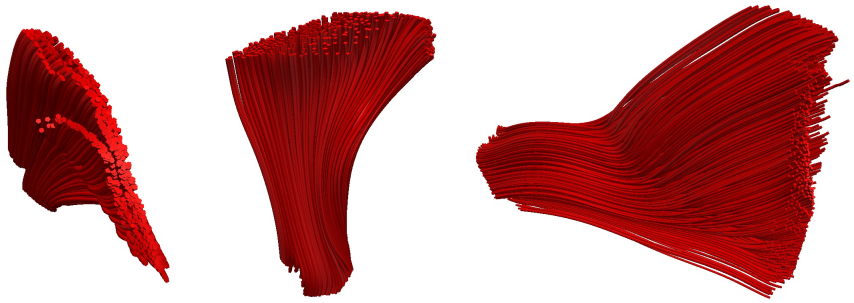

Figure 7.12: These three bundles are representative of the volumetric structure of the right corticospinal tract from three different subjects. The SGM give a well defined volume for these volumetric structures but the images suggest that the volume may not be completely filled by the fibers. This suggests that the genus of the volumetric structure may be higher than zero.

this analysis for future work.

7.8 Discussion

In this chapter we explored properties of the SGM and related them to sampling the original data in a variety of ways depending on the analysis method. In addition we showed that having a global polynomial model allows the differential geometric properties of the manifold to be derived in closed form. The SGM coefficients can also be manipulated to effect rigid body transforms. We have a new measure of connectivity and examined some of the properties of that definition.

Finally, we described using orientation statistics as a method for measuring how well the SGM fits the data. Orientation statistics was also proposed as a method to explore small-scale structure in the diffusion data that may reflect small-scale fiber structures that are not be modeled well by the SGM.

In the next chapter we implement an SGM and use some of the sampling methods we have described to analyze two white matter volumetric structures in a data set containing nominally developing control subjects versus subjects with Autism Spectrum Disorder.

Chapter 8

SGM Implementation and Experimental Evaluation

8.1 Introduction

An SGM was implemented as part of a suite of analysis programs. The suite included software to (1) build SGM models, (2) implements sampling and mapping functions, (3) outputs SGM volumes at voxel(in NIFTI format) and subvoxel levels (in Matlab mat format), (4) outputs SGM curves, $\mathcal{C}_{\parallel}, \mathcal{C}_{\perp}, \mathcal{C}_{\vdash}$ in Camino, Trackvis formats, (5) maps choice of invariants such as FA onto SGM curves and outputs each curve as a spline function ready for Functional Data Analysis, and (6) provided a wrapper around third party FDA package to do the analysis and generate reports.

For this set of experimental evaluations the software was used to extract models of white matter structures of the brain from Diffusion MRI data sets. The data sets were Diffusion Images from persons considered to be normally developing and persons diagnosed with Autism Spectrum Disorder. The normally developing persons were our control subjects and the persons with Autism Spectrum Disorder were our test subjects. The white matter volumetric structures focused on were the genu of the corpus callosum and the right corticospinal tract both of which

have been associated with Autism Spectrum Disorder. Once the SGMs were computed for each volumetric structure in each subject we use the sampling methods described in Chapter 7 to sample the Diffusion MRI data and put it in a form that allows analysis using Functional Data Analysis (FDA).

8.2 Algebraic Polynomials for Modeling SGM Surfaces

The SGM surfaces, $S_{\|}$, $S_{\perp}$ and $S_{\Vdash}$, were modeled by algebraic polynomials. The polynomials were weighed sums of global monomial basis functions. The weights are the usual coefficients of polynomial functions and constructing the model is the process of finding these weights. Letting $N = \text{order - 1}$ and $D \in \{\|, \perp, \Vdash\}$ and $\vec{x} = (x_0, x_1, x_2) \in \mathbb{R}^3$

$$S_D(\vec{x}) = \sum_{i=0}^{N}\sum_{j=0}^{N}\sum_{k=0}^{N} \beta_{(i,j,k)} x_0^i x_1^j x_2^k \tag{8.1}$$

$$\beta_{(i,j,k)} = \begin{cases} 0 & \text{, if } i+j+k > N \\ C & \text{, if } i=j=k=0 \\ \beta_{(i,j,k)} & \text{, otherwise} \end{cases}$$

where $\beta_{(i,j,k)}$ are the weights and x_0^i, x_1^j, x_2^k are the monomials raised to the powers i, j, k. N is the degree of the polynomial. For these experiments the order of all three polynomials is eight (degree seven) in three variables. The degree of the polynomial was determined empirically on a test data set. Trials on a number of different white matter structures and on a variety of data sets other than those data sets used in this evaluation appear to result in reasonable models using order eight polynomials. $\mathcal{F}(\vec{x})$ in the SGM is

$$\mathcal{F}(\vec{x}) = S_{\|}(\vec{x}) \cap S_{\perp}(\vec{x}) \cap S_{\Vdash}(\vec{x}). \tag{8.2}$$

All curves $\mathcal{C}_{\mathcal{D}}$ were defined by taking surfaces S_D pairwise ($D \in \{\|, \perp, \Vdash\}$ are space curves that, by Bézout's Theorem, may be up to 49^{th} degree polynomials. Fortunately the SGM has sufficient control of model complexity that the effective degree of the space curves is readily controlled.

The global portion of the SGM is presented in Equation 8.3. First however, a few variables are defined that make the presentation more readable. Letting $\vec{U}_{\|}$ be the unit vector $\frac{\nabla S_{\|}(\vec{x})}{\|\nabla S_{\|}(\vec{x})\|}$, $\vec{U}_{\perp}$ the unit vector $\frac{\nabla S_{\perp}(\vec{x})}{\|\nabla S_{\perp}(\vec{x})\|}$, and $\vec{U}_{\Vdash}$ the unit vector $\frac{\nabla S_{\Vdash}(\vec{x})}{\|\nabla S_{\Vdash}(\vec{x})\|}$. Let $H(S_{\|})$, $H(S_{\perp})$ and $H(S_{\Vdash})$ be the Hessian matrices with rows i and columns j. Then

$$
\begin{aligned}
\mathfrak{F}(\vec{x}) = \min_{x \in \mathfrak{M}} \Bigg\{ & \lambda_1 \|\vec{U}_{\|} \cdot \vec{\varepsilon}_{\|}\|_{L_\infty} + \lambda_2 \left[\arg\min_{x \in \mathfrak{M}} \| d_{\mathbb{G}}(S_{\|}^{M(\vec{\mu})}(\vec{x})|_{L_1} \right] \\
& + \lambda_3 \|\vec{U}_{\perp} \cdot \vec{\varepsilon}_{\perp}\|_{L_\infty} + \lambda_4 \left[\arg\min_{x \in \mathfrak{M}} \| d_{\mathbb{G}}(S_{\perp}^{M(\vec{\mu})}(\vec{x})\|_{L_1} \right] \\
& + \lambda_5 \|\vec{U}_{\Vdash} \cdot \vec{\varepsilon}_{\Vdash}\|_{L_\infty} + \lambda_6 \left[\arg\min_{x \in \mathfrak{M}} \| d_{\mathbb{G}}(S_{\Vdash}^{M(\vec{\mu})}(\vec{x})\|_{L1} \right] \\
& + \lambda_7 \left\| 1 - (\vec{U}_{\|} \cdot (\vec{U}_{\perp} \times \vec{U}_{\Vdash})) \right\|_{L_2} \\
& + \lambda_8 \| \sum_{i,j} H(S_{\|}(\vec{x}))_{i,j} \|_{L_2} \\
& + \lambda_9 \| \sum_{i,j} H(S_{\perp}(\vec{x}))_{i,j} \|_{L_2} \\
& + \lambda_{10} \| \sum_{i,j} H(S_{\Vdash}(\vec{x}))_{i,j} \|_{L_2} \Bigg\}.
\end{aligned}
\tag{8.3}
$$

8.2.1 Modeling Local Structure with the SGM Support Vector Machine

The global polynomial portion of the model required the data to lie on a single unknown manifold. Conversely, as indicated in Equation 8.4, the SGM Support Vector Machine (SSVM) needs the SGM global function to help it "decide" if the

data lie on a manifold that is consistent with the global polynomial. Therefore some indication is needed of what data are initially part of the manifold that we want to model. This is an example of semi-supervised manifold learning. In addition to giving some examples of where to start, the SSVM needs a description of how the data should be considered similar to one another. Since the SSVM only selects which data to add "near" other data and "near" the SGM global function the similarities need to be evaluated locally. Given those requirements and the mathematical requirements for the SSVM, similarities are defined by kernel functions that are local or compact. Scaling each kernel separately as described below allows controlling how similar data points should be. In this way the kernels "find" the local structure of the manifold guided by the SGM global function.

Recalling Equation 6.17, let

$$\langle \mathbf{w}, \Phi(x_i) \rangle = \mathbb{K}(\bullet, \bullet_i) \tag{8.4}$$

$$\min_{\mathbf{w} \in \mathfrak{H}, \xi^{(*)} \in \mathbb{R}^m} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{m\nu}\sum_{i=1}^{m}(\xi_i^* + \xi_i), \tag{8.5}$$

$$\text{subject to} \begin{cases} -\xi_i \leqslant \mathbb{K}(\bullet, \bullet_i) + \lambda\mathcal{F}(x_i) \leqslant \tau + \xi_i^* \\ \xi^{(*)} \geqslant 0, \end{cases}$$

Each similarity kernel is a compact radial basis function. The radial basis functions chosen for this study were Gaussian inside a scaled interval and zero outside of that interval. Recall that the three eigenvectors of the Diffusion Tensor **D** are represented by $\varepsilon_{\parallel}, \varepsilon_{\perp}, \varepsilon_{\Vdash}$ and the kernels are given by:

$$K_d(x, x_i) = e^{(-\gamma_1 * (\|x - x_i\|_2))} \tag{8.6}$$

$$K_{\parallel}(\vec{\varepsilon}_{\parallel}, \vec{\varepsilon}_{\parallel i}) = e^{(-\gamma_2 * (1 - |\langle \vec{\varepsilon}_{\parallel}, \vec{\varepsilon}_{\parallel i} \rangle|))} \tag{8.7}$$

$$K_{\perp}(\vec{\varepsilon}_{\perp}, \vec{\varepsilon}_{\perp i}) = e^{(-\gamma_3 * (1 - |\langle \vec{\varepsilon}_{\perp}, \vec{\varepsilon}_{\perp i} \rangle|))} \tag{8.8}$$

$$K_{FA}(x, x_i) = e^{(-\gamma_4 * (1 - |FA(\vec{x}) - FA(\vec{x}_i)|))}. \tag{8.9}$$

$K_d(*)$ controls the size of the neighborhood of the model. $K_{\|}(*)$ is a similarity metric for the difference in orientation between the $\varepsilon_{\|}$ eigenvectors. $K_{\perp}(*)$ is a similarity metric for the difference in orientation between the $\varepsilon_{\perp}$ eigenvectors. Finally, $K_{FA}(*)$ is a similarity metric for the diffusivity properties of the diffusion MRI data. For these experiments $K_{FA}(*)$ is the fractional anisotropy at each data point. These kernels were combined into the kernel

$$\mathbb{K}(\bullet, \bullet_i) = \prod_{j=1}^{4} K_j(\bullet, \bullet_i), \tag{8.10}$$

where $j \in \{d, \|, \perp, FA\}$. Note: for these experiments our similarity kernels were measuring the difference in orientation between tensors. Since the eigenvectors of the tensor are orthogonal, the orientation of the tensor in $\mathbb{R}^3$ is fully determined by any two of them.

8.3 SGM and Functional Data Analysis

When performing an analysis it is important that correlations between data points be taken into account. For example, voxel-based methods assume that all voxels are statistically independent [10]. However, as outlined in Section 2.2, white matter is composed of continuous fibers and therefore measurements taken along the fibers are likely to be highly correlated. Voxel-based methods also apply Gaussian smoothing kernels to the diffusion data at each voxel. This has the effect of reducing the effects of imperfect registration as well as increasing the statistical power for subsequent analysis. Unfortunately, as pointed out previously [135–137], if the data are already correlated in a given direction this kind of smoothing introduces spurious additional correlations. A method of statistical analysis that takes the kind of correlation in our data into account is Functional Data Analysis (FDA).

8.3.1 Choosing Sampling Surfaces

There are formal methods for picking corresponding sampling surfaces across subjects but those remain part of our future work. Studies [60, 67] suggest that analysis of white matter along the direction of the fibers is not very sensitive to cross-subject alignment. A preliminary look at curve lengths in the data showed that the curves in the structures modeled were similar in length. So, for the purposes of this analysis, a simple approach is taken. Assuming good extracted volumes by SGMs, a correspondence between structures is assumed to be the Median Curve (MC) of each structure. An MC, $\mathcal{C}_{\parallel}^{M}$, is a curve determined by the intersection of two median surfaces, $S_{\perp}^{M} \bigcap S_{\shortparallel\vdash}^{M}$. MCs were determined for given structures in each subject. Then, each MC is partitioned into 200 evenly-spaced, non-overlapping segments. Orthogonal sampling surfaces, $S_{\parallel}$ were determined for each MC, for each subject. The two sampling methods described below use these surfaces to map FA values to sampling curves.

8.3.2 Functional Mapping

The simplest form of Functional Data Analysis [135, 136] assumes that the data can be organized into a number of single parameter curves [135–137]. It is straight forward to use an SGM to sample the voxel space of a white matter structure and produce curves comprised of FA values.

For example using the following equation:

$$\begin{aligned} \mathcal{C}_{\parallel}(t,s) &= \mathfrak{M}(\nu_{\parallel}(t), \nu'_{\perp}, \nu_{\shortparallel\vdash}(s)) \\ &= S_{\parallel}(\vec{x},t) \cap S'_{\perp}(\vec{x}) \cap S_{\shortparallel\vdash}(\vec{x},s). \end{aligned} \tag{8.11}$$

an initial $\vec{x}$ defines a particular curve t that is the intersection $S_{\parallel}(\vec{x},t) \cap S'_{\perp}(\vec{x})$. s indexes along the curve. Then the mapping function $\mathcal{M}$ takes the manifold curves and FA data and maps them to a curve with the discrete FA values. More formally, $\mathcal{M}(\mathcal{C}_{\parallel}(t,s), FA(t,s)) \mapsto \mathcal{C}_{FA}(t)$.

The next step is to estimate a smooth function $\mathfrak{C}$ of the FA values using the $\mathcal{C}_{FA}$ curves.

8.3.3 Functional Smoothing

The next step in preparing the data for analysis is smoothing. Previous work [133, 136] suggests smoothing the data using a penalized least-squares method:

$$\mathcal{S}(t,\mathfrak{C}(t),\lambda) = \sum_{j} (\mathcal{C}_{FA}(t_j) - \mathfrak{C}(t_j))^2 + \lambda \int \left(\frac{d^n \mathfrak{C}(t)}{dt^n} \right)^2 dt \tag{8.12}$$

where λ is a parameter that controls the amount of roughness allowed in the solution. A choice of basis functions needs to be made as well:

$$\mathfrak{C}(t) = \sum_{k=1}^{K} \beta_k \phi_k(t) \tag{8.13}$$

where $\phi_k(t)$ is the k^{th} basis function, β_k is the k^{th} weight and K is the number of basis functions. Given that the data are effectively height functions of FA values in the range of $[0,1]$, B-splines were a reasonable choice for the basis functions. With these choices made, the specific description for the functional data can now be made.

As described in Section 8.3.1, the $\mathfrak{C}_{FA}(t)$ were sampled at 200 points and the number of basis functions K were chosen as 190. The order of the B-splines was set to 6. n, the order of the derivative used in the smoothing function, was set to 4. The analysis showed little sensitivity to these choices. A more important issue is the choice of the smoothing parameter λ because the results of the analysis are somewhat sensitive to the choice of λ. The value of λ was determined using a combination of generalized cross validation [42] and empirical observation.

8.3.4 Functional Registration

Determining the $\mathcal{C}_{\|}(t)$ curves within a single subject is well defined using the SGM. However, when analyzing the data between subjects using the approximation methods described here, exact correspondence between orthogonal sampling is uncertain. Functional registration or curve registration [137] compensates for this uncertainty as well as providing a method to evaluate the sensitivity of the analysis to this sampling method. The registration step used here consists of calculating a functional mean curve using all of the data from both the control and autism groups for a given structure. Then, data curves were registered using methods described in [137]. Registration appears to reduce variance somewhat but the results were identical with and without registration. One reason this might be correct is the observed relative insensitivity of alignment errors along white matter fibers noted in Section 8.3.1.

8.4 Experimental Design

Differences in white matter structure between persons diagnosed with Autism Spectrum Disorder have been noted in a number of studies. Two white matter structures that have been shown to exhibit such differences were chosen for modeling and analysis: (1) a portion of the genu of the corpus callosum, and (2) the right corticospinal tract. The genu has interesting internal structure because it has a curved global shape and the cingulum is closely associated with it. An SGM that models the genu exclusively requires that a few voxels specific to the genu be indicated. In principal these additional voxels weren't strictly needed but using them greatly reduces the time need to compute an SGM. Large ROIs were manually specified and used in post-processing to indicate which structure the model should keep.

The sampling surfaces were chosen based on the method described in Section 8.3.1. Median Curve(MC)s were determined for given structures in each subject. Then, each MC is partitioned into 200 evenly-spaced, non-overlapping segments.

Orthogonal sampling surfaces, $S_{\parallel}$ were determined for each MC, for each subject.

Sampling Method 1. *Calculate the Median Curve* $\mathcal{C}_{\parallel}^{M}$ *and the orthogonal sampling surface* $S_{\parallel}^{i}$ $i = 1...200$. *Find the maximum FA value on each surface. Then* $\mathcal{M}(\mathcal{C}_{\parallel}(t), max(FA)) \mapsto \mathcal{C}_{max(FA)}(t)$. *The result is a single curve* $\mathcal{C}_{max(FA)}(t)$ *representing each structure in each subject.*

Sampling Method 2. *Calculate the Median Curve* $\mathcal{C}_{\parallel}^{M}$ *and the orthogonal sampling surface* $S_{\parallel}^{i}$ $i = 1...200$. *Use the SGM for each structure to generate a large number* N *of curves* $\mathcal{C}_{\parallel}$ *randomly distributed in the structure volume. Then construct the curves that are the intersection of each* $\mathcal{C}_{\parallel}$ *curve with the orthogonal sampling surfaces. For each intersection point find the FA value at that point. Then* $\mathcal{M}(\mathcal{C}_{\parallel}^{n}(t), FA) \mapsto \mathcal{C}_{FA}^{n}(t)$ *where* $n \in \{1, ..., N\}$. *The output is sets of* N $\mathcal{C}_{FA}$ *curves. Each set of* N *curves samples the volume of a single structure.*

$\mathcal{C}_{FA}(t)$ curves were produced for each sampling method. First derivative curves ($\mathcal{C}'_{FA}(t)$) and second derivative curves ($\mathcal{C}''_{FA}(t)$) were calculated as well. This allows the exploration of the distribution of FA values, the rate of change in the FA values and the "acceleration" in the rate of change.

8.4.1 Data

The data used in this evaluation were provided by the Waisman Laboratory for Brain Imaging and Behavior University of Wisconsin-Madison. After initial registration and tensor reorientation the data sets were resampled to an isotropic resolution of 1mm x 1mm x 1mm. The tensor and FA maps were then calculated using the Camino software package [39]. Finally all of the visualizations shown were a combination of Trackvis [182], fslview [150] and Matlab [110]. Functional Data Analysis was performed using Matlab code described in [134]. In the next section an analysis of the genu of the corpus callosum is presented.

8.5 FA Maps of the Input Data

Figures 8.1 and 8.2 were FA maps of sagittal view slices through the midline of the brain for nine subjects from the control group and nine subjects from the autism group. The corpus callosum (cc) is the large, white horizontal structure. Notice that there is considerable variation in the shape of the cc near the genu. The usual methods of analysis would require further registration or using large smoothing kernels to do analysis whereas our analysis proceeds without further registration. The FDA analysis focuses on the genu of the cc located on the right end of the cc in each image, and labeled in red. To further emphasize that these extracted volumes were only roughly aligned, see Figures 8.3 and 8.4. Figure 8.3 show two views of the overlapping genu volumes of the control subjects and Figure 8.4 depicts two views of the overlapping genu volumes of the autism subjects. Perfectly aligned volumes would result in monochrome images.

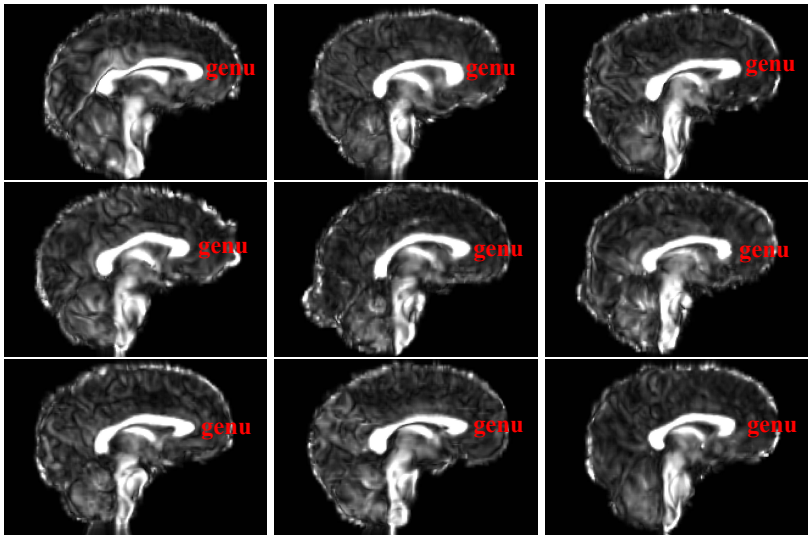

Figure 8.1: Sagittal view slices through the midline of the brain for nine control subjects. Notice the variation in shape of the region labeled genu.

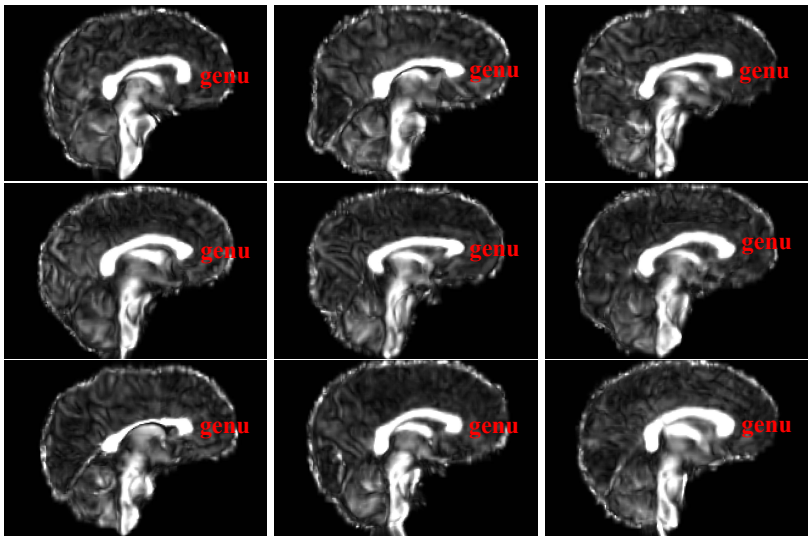

Figure 8.2: Sagittal view slices through the midline of the brain for nine autism subjects. Notice the variation in shape of the region labeled genu.

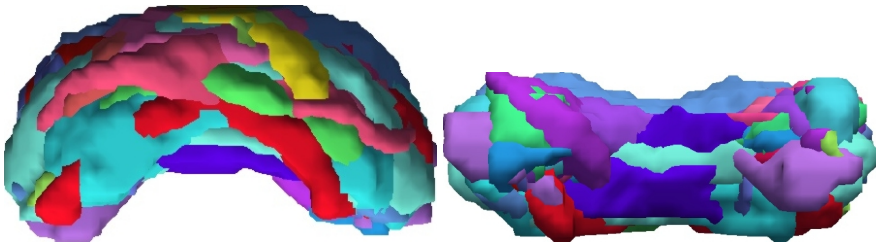

Figure 8.3: Two views of the overlapping volumes of extracted portions of the genu for the control group. Each subject's volume is a different color. A perfect registration would show only one color.

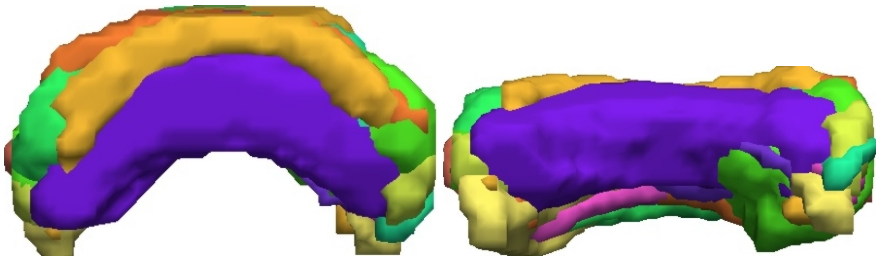

Figure 8.4: Two views of the overlapping volumes of extracted portions of the genu for the autism group. Each subject's volume is a different color. A perfect registration would show only one color.

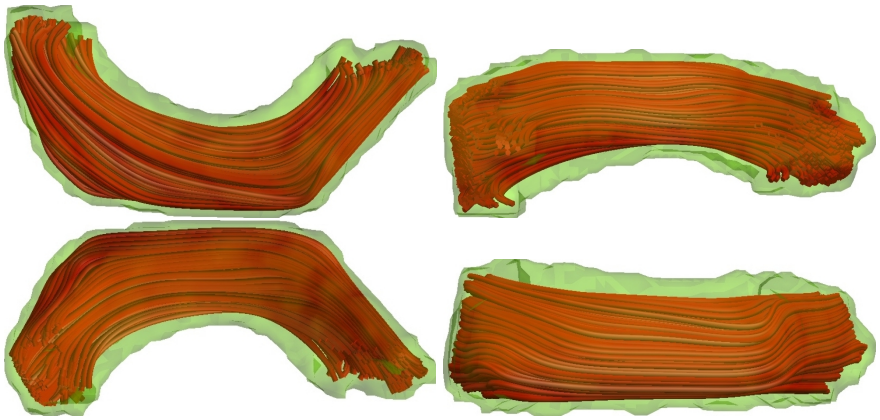

Figure 8.5: Four views of an SGM model of the genu of the cc of a control subject. Green shows the voxels of the SGM volume, and SGM curves through the volume are shown in red.

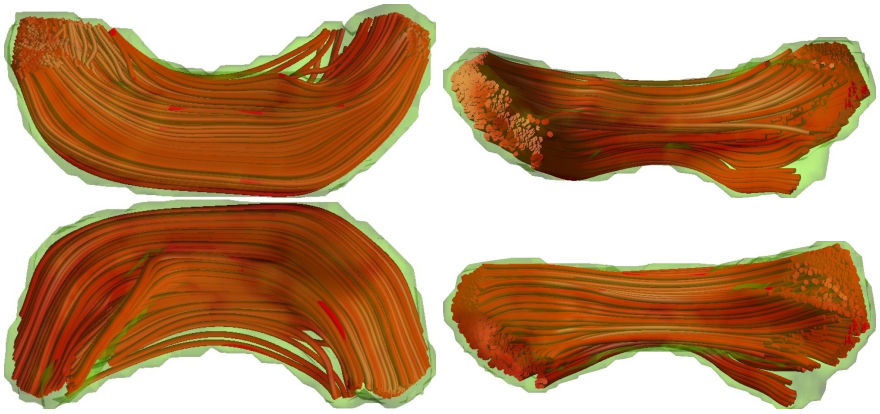

Figure 8.6: Four views of an SGM model of the genu of the cc of a control subject. Green shows the voxels overlaying the SGM volume, and SGM curves through the volume are shown in red.

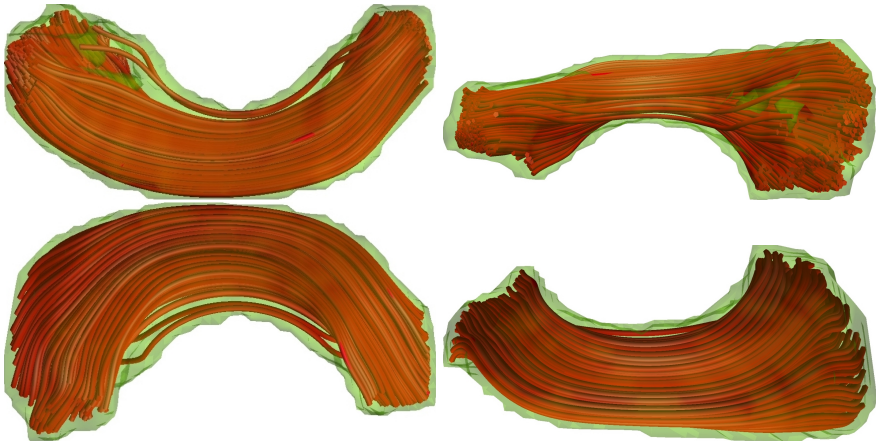

Figure 8.7: Four views of an SGM model of the genu of the cc of an autism subject. Green shows the voxels overlaying the SGM volume, and SGM curves through the volume are shown in red.

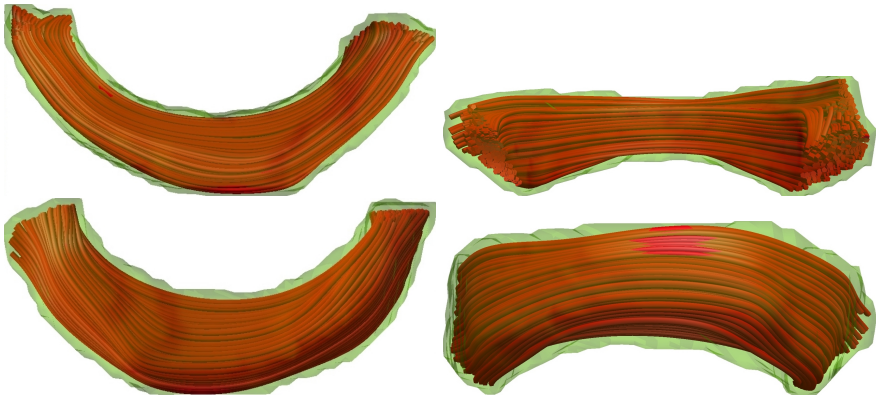

Figure 8.8: Four views of an SGM model of the genu of the cc of an autism subject. Green shows the voxels overlaying the SGM volume, and SGM curves through the volume are shown in red.

8.6 Sampling the Genu

In the course of analyzing a data set it is important and useful to visualize elements of the data set. Once an SGM has been computed for the genu in each subject, the overall organization of the tracts modeled by the SGM can be viewed by sampling the volume with sets of $\mathcal{C}_{\|}$. Figures 8.5 and 8.6 show a set of $\mathcal{C}_{\|}$ curves traced through the body of the genu of a control subject. The shaded green areas are the voxels representing the volume extracted by the SGM. Likewise, Figures 8.7 and 8.8 represent the organization of the white matter fibers by $\mathcal{C}_{\|}$ curves traced through the body of the genu of a subject with autism.

Figure 8.9 shows and axial and coronal views of the genu and the orthogonal SGM curves. This visualization gives a sense of the global organization of the white matter fibers in terms of smooth functions of the three sets of eigenvectors of the Diffusion Tensor field.

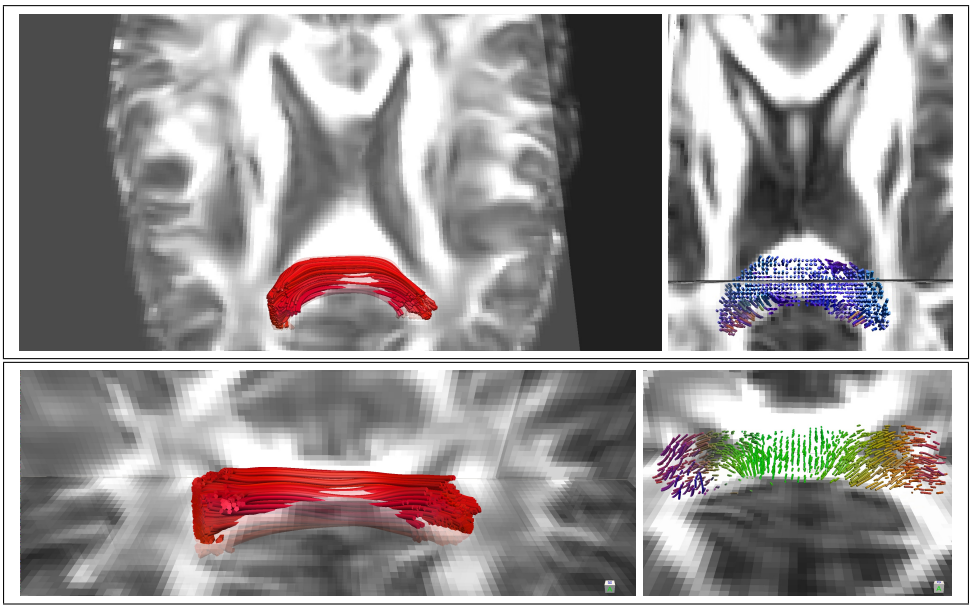

Figure 8.9: (top left) $\mathcal{C}_{\parallel}$ curves superimposed over the genu on an axial view FA map. (top right) $\mathcal{C}_{\parallel\vdash}$ curves superimposed over the genu on an axial view FA map. (bottom left) $\mathcal{C}_{\parallel}$ curves superimposed over the genu on a coronal view FA map. (bottom right) $\mathcal{C}_{\perp}$ curves superimposed over the genu on a coronal view FA map.

8.7 Analyzing the Genu – Sampling Method 1

The genu of the corpus callosum is one of the regions that has been shown to exhibit differences in white matter composition between people with normal development and people with Autism Spectrum Disorder [5, 177]. Sampling Method 1, as depicted in Figure 8.10, is used to generate the functional data for this analysis. This results in sixteen smoothed max(FA) curves representing the control group and thirteen smoothed max(FA) curves representing the autism group. Figure 8.11 shows plots of these FA curves. The curves were registered to the mean curve (see Figure 8.12) and the results plotted in Figure 8.13. T-tests compared the curves generated using Sampling Method 1. No significant differences were found between control and autism groups when simply comparing the distribution of max(FA) values along the curves. However, the $F'(t)$ and $F''(t)$ curves do show statistically significant differences at the .01 significance level on the right side of genu. The biological interpretation of the these differences remain to be determined.

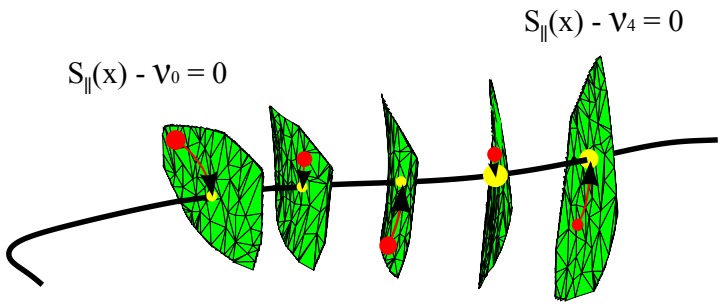

Figure 8.10: Genu of the corpus callosum: Sampling Method 1 – Maximum FA from each orthogonal surface projected onto $\mathcal{C}_{\|}^{M}$.

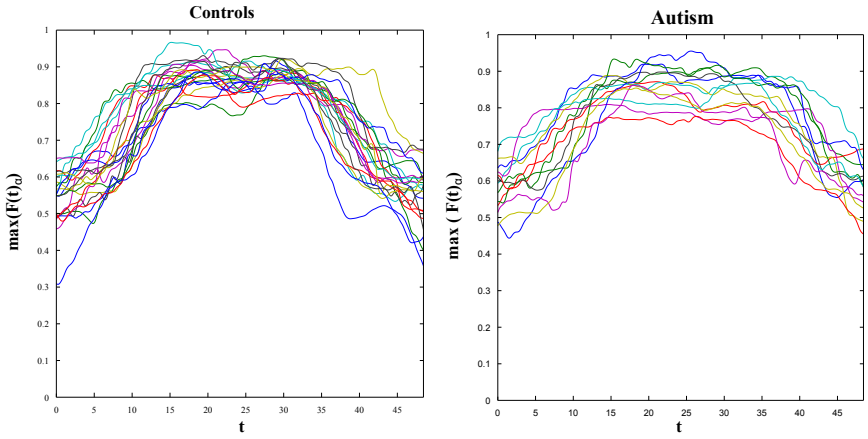

Figure 8.11: Unregistered Maximum FA $\mathsf{F}(\mathsf{t})$ curves of the genu for the control subjects on the left and the autism subjects on the right.

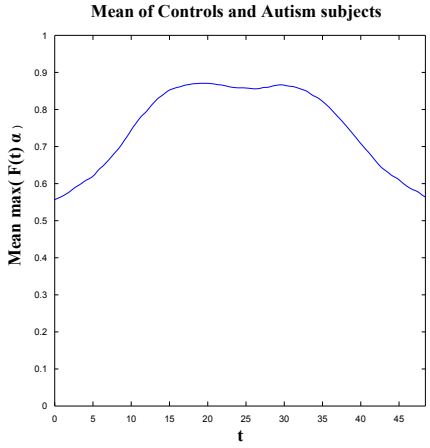

Figure 8.12: Genu of the corpus callosum: Mean Curve of the maximum FA curves projected onto $\mathcal{C}^{\mathsf{M}}$ curve. This is the reference curve for registering both the control and test data.

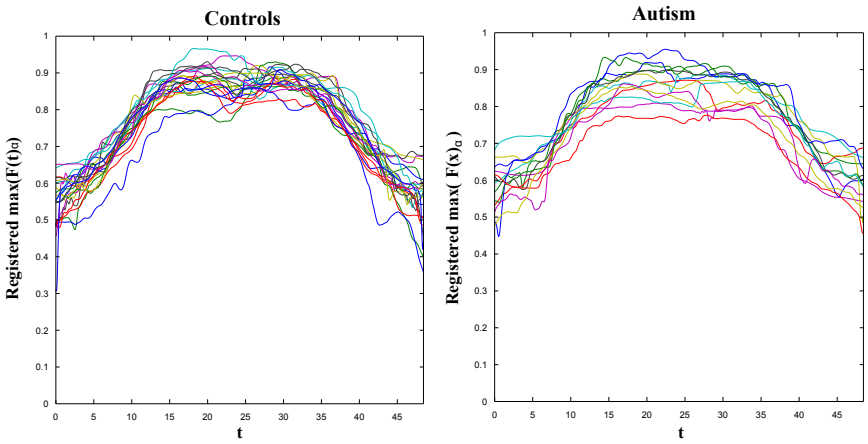

Figure 8.13: Maximum FA $F(t)$ curves of the genu (left) control and (right) autism subjects.

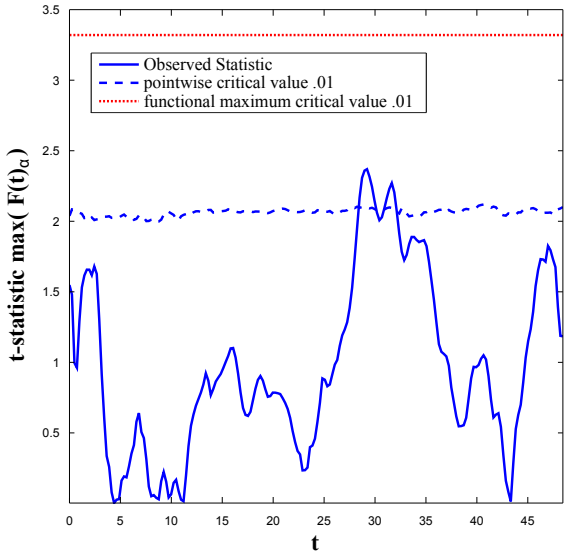

Figure 8.14: Genu: Maximum FA values Functional t-test controls vs autism for $F(t)$. At around $t = 30$ the pointwise t-statistic reached significance however that statistic doesn't take correlation of the data along the curves into account. For the functional t-test there is no statistically significant difference between the control and autism curves.

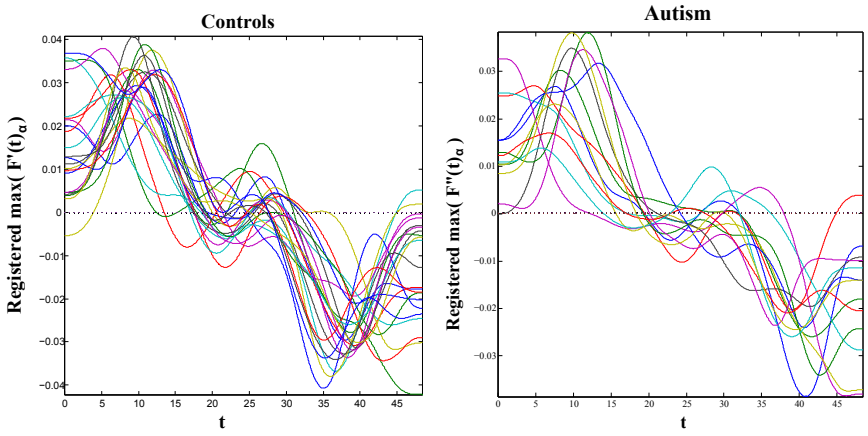

Figure 8.15: Genu: Maximum FA $F'(t)$ indicates the rate of change with respect to t in the FA curves. (left) control subjects (right) autism subjects.

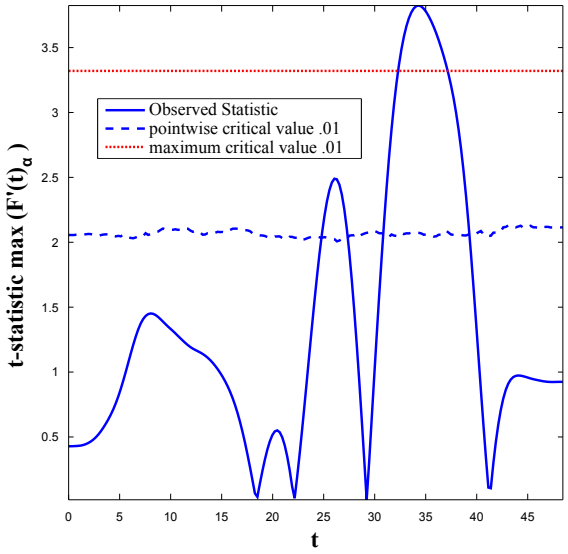

Figure 8.16: Genu: Maximum FA values Functional t-test controls vs autism for $F'(t)$. At $t = 25$ and $t = 35$ the pointwise t-statistic reached significance however that statistic doesn't take correlation of the data along the curves into account. For the functional t-test there is a statistically significant difference at the .01 significance level, between the control and autism curves.

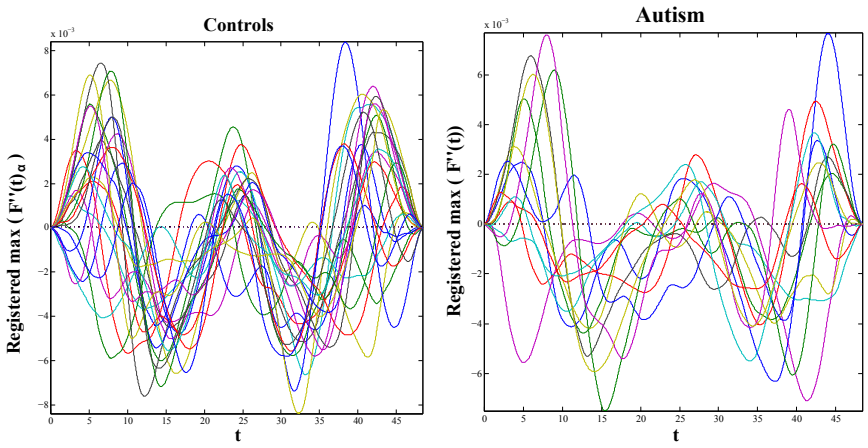

Figure 8.17: Genu: Maximum FA $\mathrm{F}''(\mathrm{t})$ indicates the "acceleration" in the change with respect to t in the FA curves. (left) control subjects, (right) autism subjects.

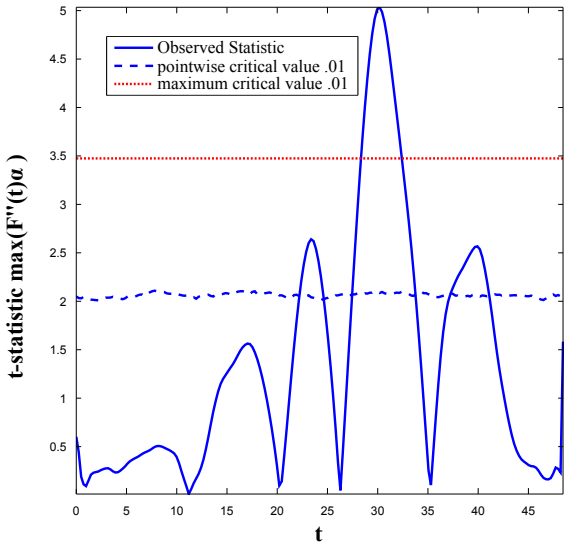

Figure 8.18: Genu: Maximum FA values Functional t-test controls vs autism for $\mathrm{F}''(\mathrm{t})$. At around $\mathrm{t} = 23, 30, 40$ the point-wise t-statistic reached significance however that statistic doesn't take correlation of the data along the curves into account. For the functional t-test there is a statistically significant difference at the .01 level between the control and autism curves.

8.8 Analyzing the Genu – Sampling Method 2

Sampling Method 2 as depicted in Figure 8.19 was used to generate the functional data for this analysis. For this analysis the curves sampled from the control group were bulked together and sampled randomly to get 1000 curves that represent the entire control group of subjects. The curves from the autism group was also bulked and sampled randomly to get 1000 curves that represent the entire autism group of subjects. This results in smoothed FA curves representing the control group and smoothed FA curves representing the autism group. The curves were registered to the mean curve and the results plotted in Figure 8.20. T-tests compared the curves generated using Sampling Method 2. The $F'(t)$ and $F''(t)$ curves were generated by taking the first and second derivatives of the $F(t)$ curve, respectively. Most of $F(t)$ showed significant differences between the control and autism groups when simply comparing the distribution of FA values along the curves. $F'(t)$ showed statistically significant differences at the .01 significance level on most of the curve. $F''(t)$ showed statistically significant differences at the .01 significance level at the left and right ends and just right of the center of the genu. The biological interpretation of these differences remains to be determined.

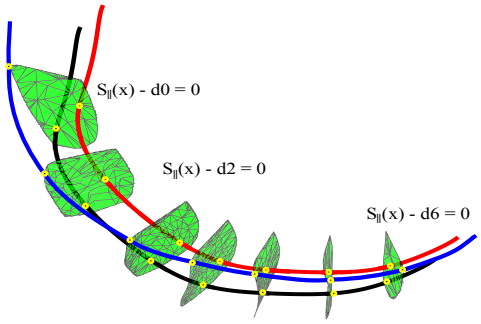

Figure 8.19: Sampling multiple curves through the genu using a common set of orthogonal surfaces.

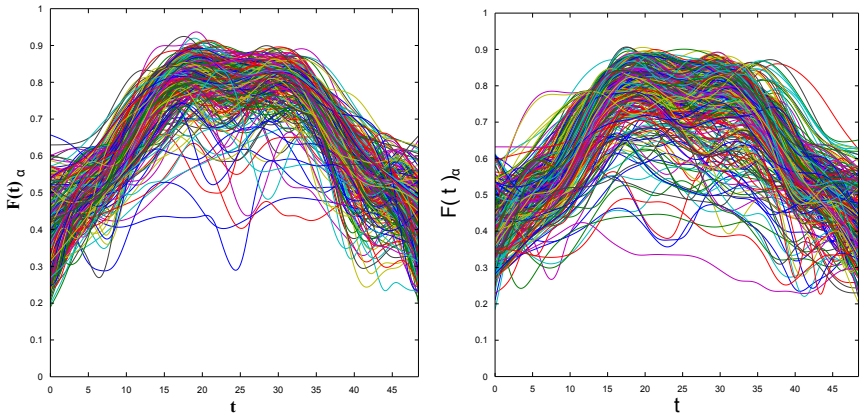

Figure 8.20: Genu – (left) control subjects. (right) autism subjects. Sampling method 2. $F(t)$ curves.

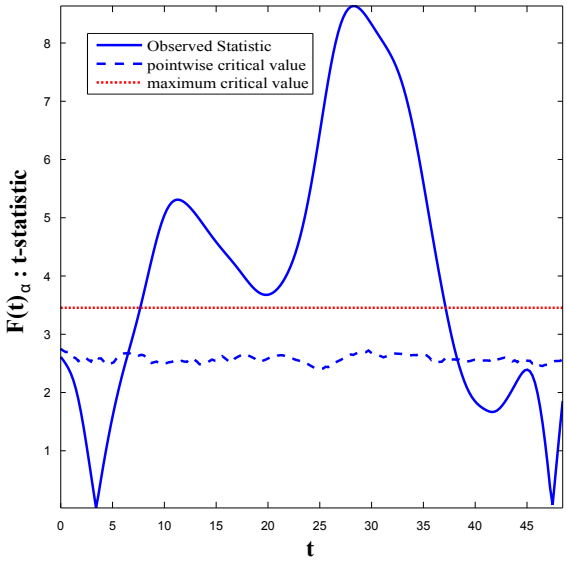

Figure 8.21: Genu fully sampled volume – Functional t-test controls vs autism for $F(t)$. Most of the curve showed statistically significant differences at the .01 significance level.

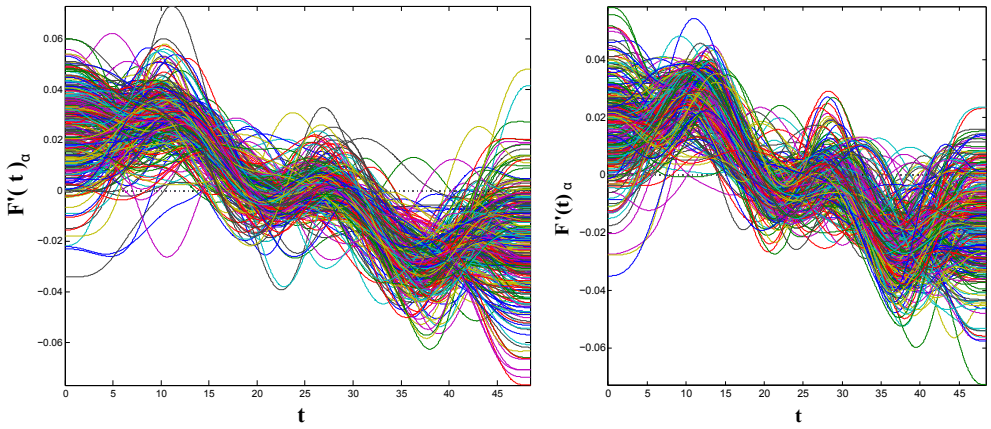

Figure 8.22: Genu – (left) control subjects. (right) autism subjects: Sampling method. 2. $F'(t)$ curves.

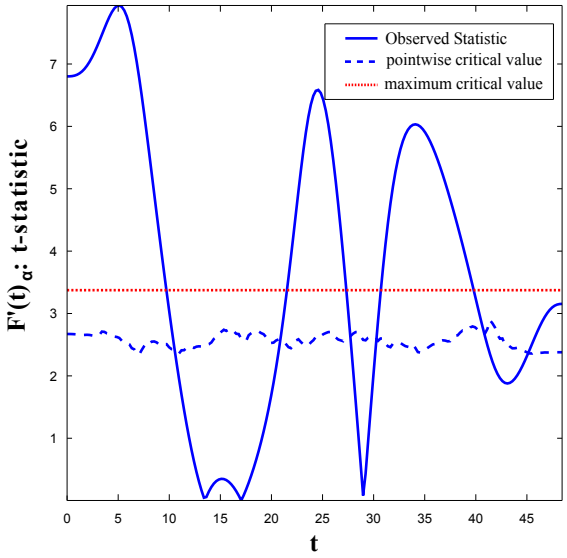

Figure 8.23: Genu fully sampled volume – Functional t-test controls vs autism for $F'(t)$. Most of the curve showed statistically significant differences at the .01 significance level.

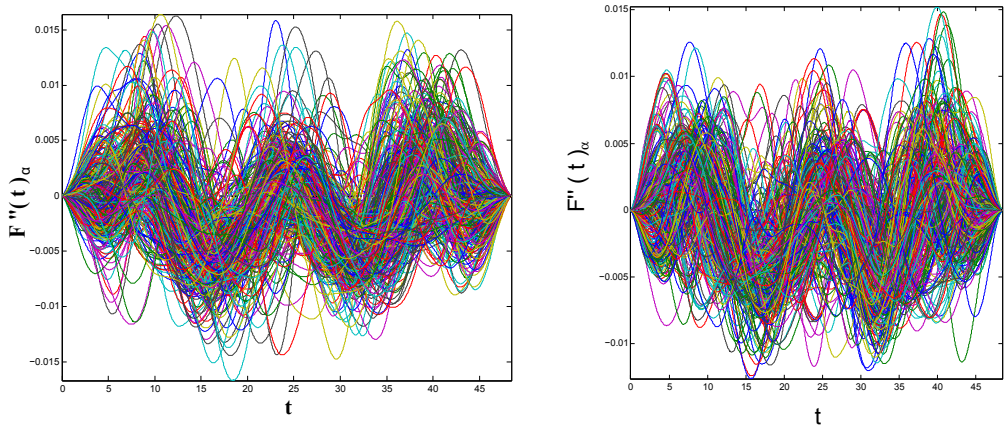

Figure 8.24: Genu – (left) control subjects. (right) autism subjects: Sampling method. 2 for $\mathrm{F}''(\mathrm{t})$ curves.

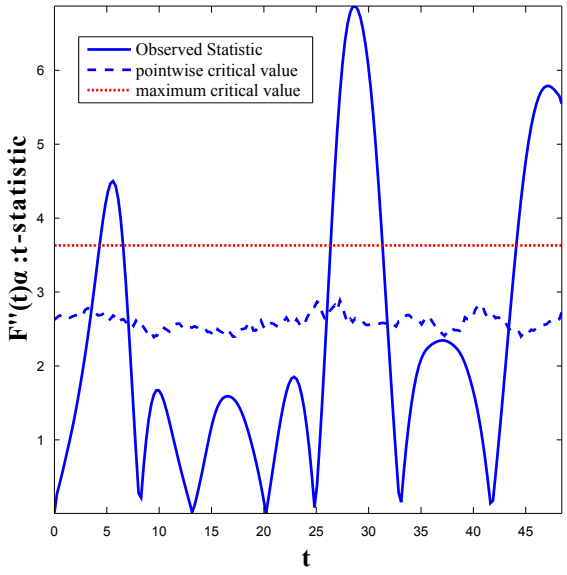

Figure 8.25: Genu fully sampled volume – Functional t-test controls vs autism for $\mathrm{F}''(\mathrm{t})$. Three areas along the curve show statistically significant differences at the .01 significance level. They are the left and right ends of the genu and just right of the center of the genu.

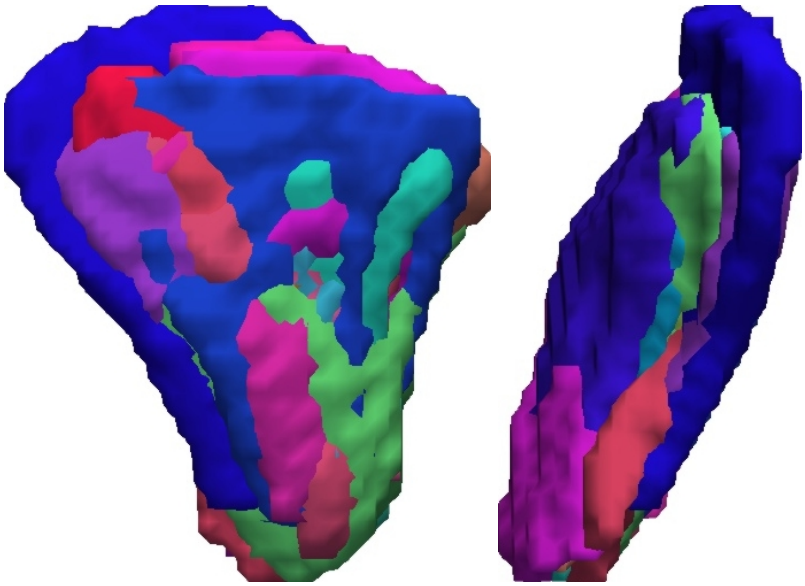

Figure 8.26: Corticospinal registration. Two views of the overlapping volumes of SGM volumes of portions of the corticospinal tract for the control group. A perfect registration would show only on color.

8.9 Sampling the Corticospinal Tract

Once an SGM has been computed for the right side of the corticospinal tract in each subject, the overall organization of the tracts modeled by the SGM can be viewed by sampling the volume with sets of $\mathcal{C}_{\|}$. Figures 8.5 and 8.6 show a set of $\mathcal{C}_{\|}$ curves traced through the body of the right corticospinal tract of a control subject. The shaded green areas are the voxels representing the volume extracted by the SGM. These figures were generated by data from the control group. Likewise, Figures 8.7 and 8.8 represent the organization of the white matter fibers by $\mathcal{C}_{\|}$ curves.

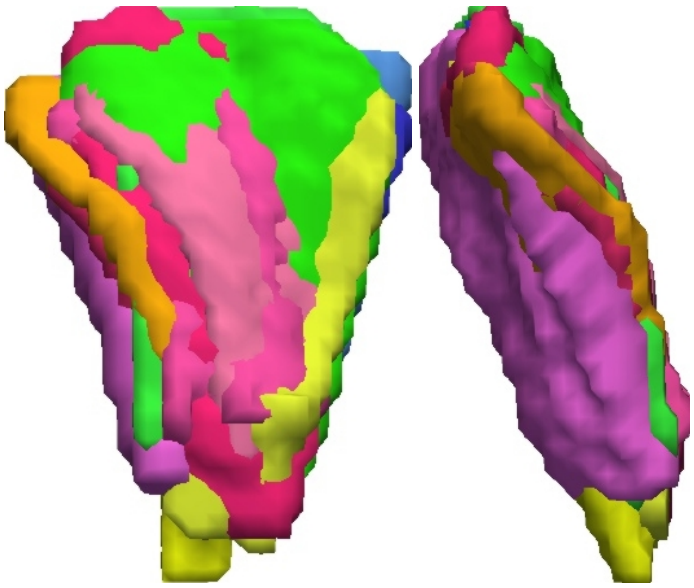

Figure 8.27: Corticospinal registration. Two views of the overlapping SGM volumes of portions of the corticospinal tract for the autism group. Perfect registration would show only one color.

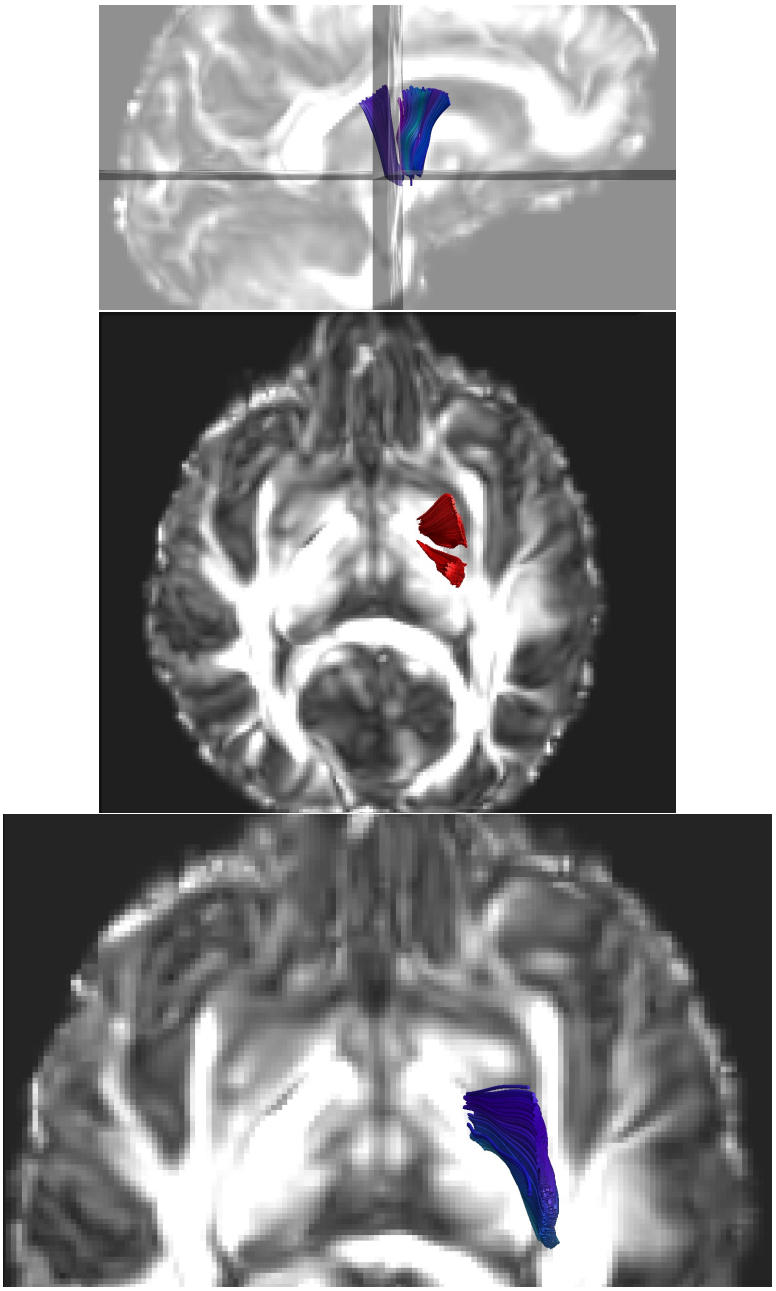

Figure 8.28: Corticospinal track: The top image is a combination of cortico-pontine and cortico-bulbar fibers extracted simultaneously by an SGM superimposed onto sagittal and axial view FA maps, respectively. The bottom image is the cortico-pontine tract superimposed on an axial view FA map.

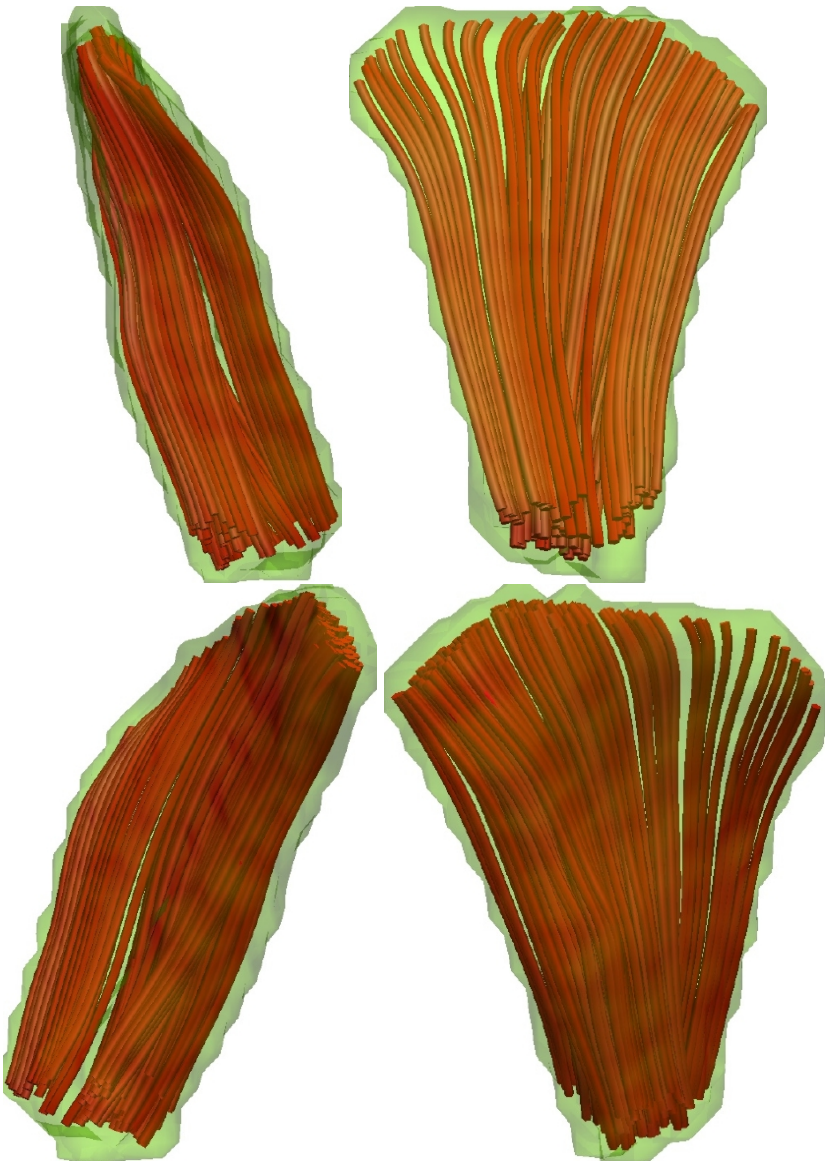

Figure 8.29: Four views of an SGM volume and modeling of the right corticospinal tract of a control subject. Green shows voxels overlaying the SGM volume and SGM $\mathcal{C}_{\|}$ curves through the volume are shown in red.

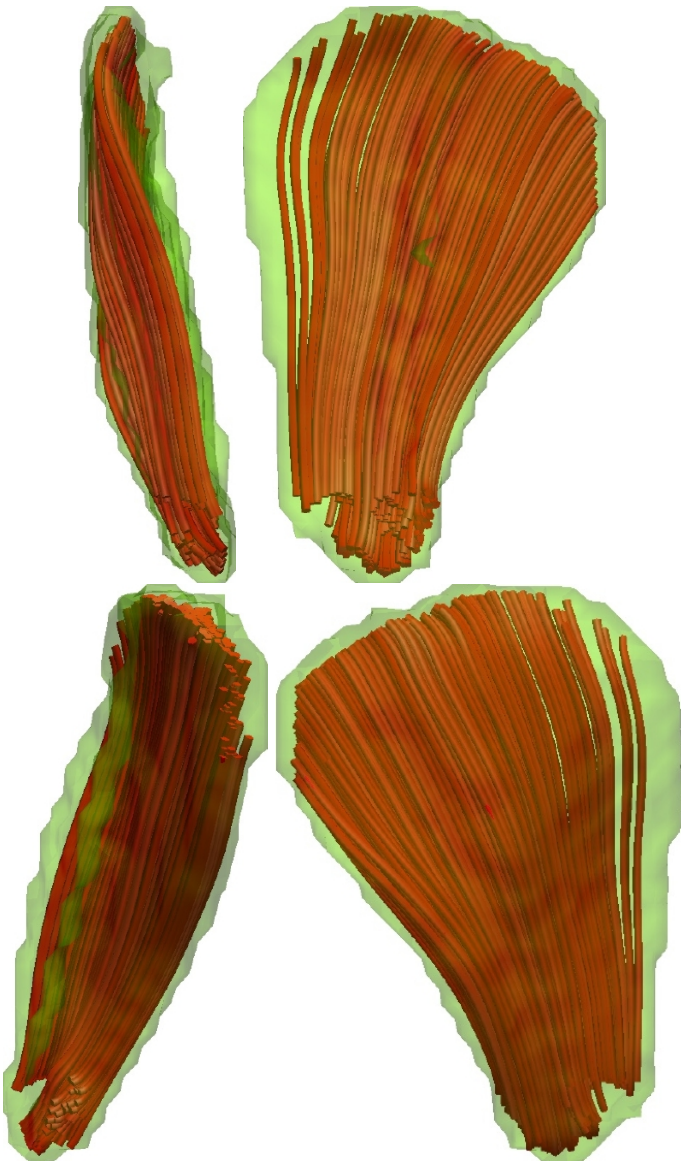

Figure 8.30: Four views of an SGM volume and modeling of the right corticospinal tract of a control subject. Green shows voxels overlaying of the SGM volume and SGM $\mathcal{C}_{\|}$ curves through the volume are shown in red.

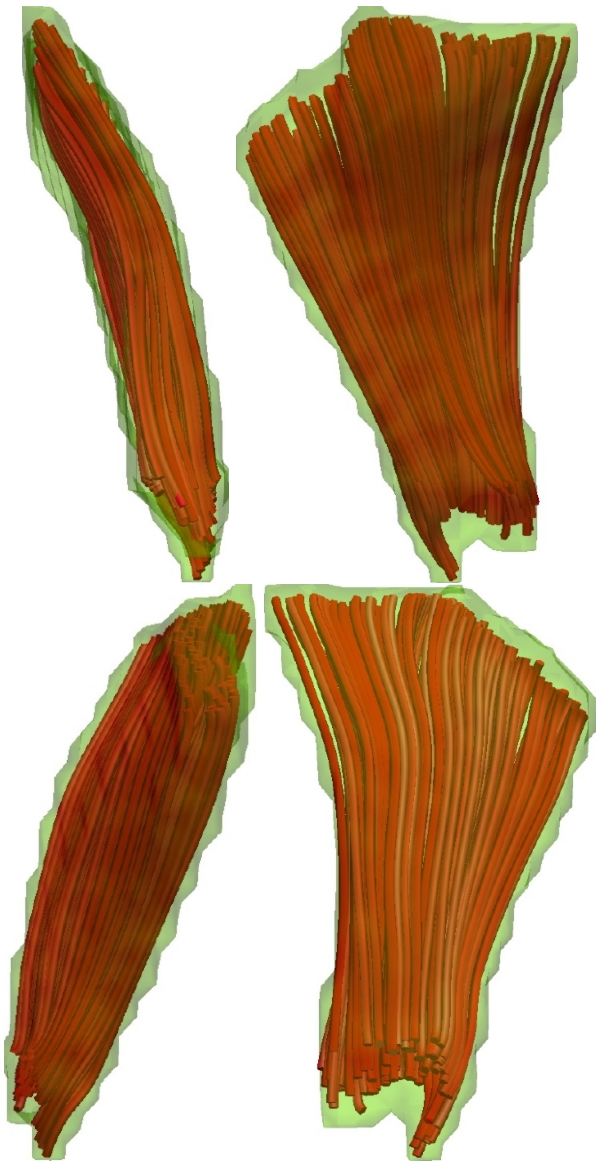

Figure 8.31: Four views of an SGM volume and modeling of the right corticospinal tract of an autism subject. Green shows voxels overlaying the SGM volume and SGM $\mathcal{C}_{\|}$ curves through the volume are shown in red.

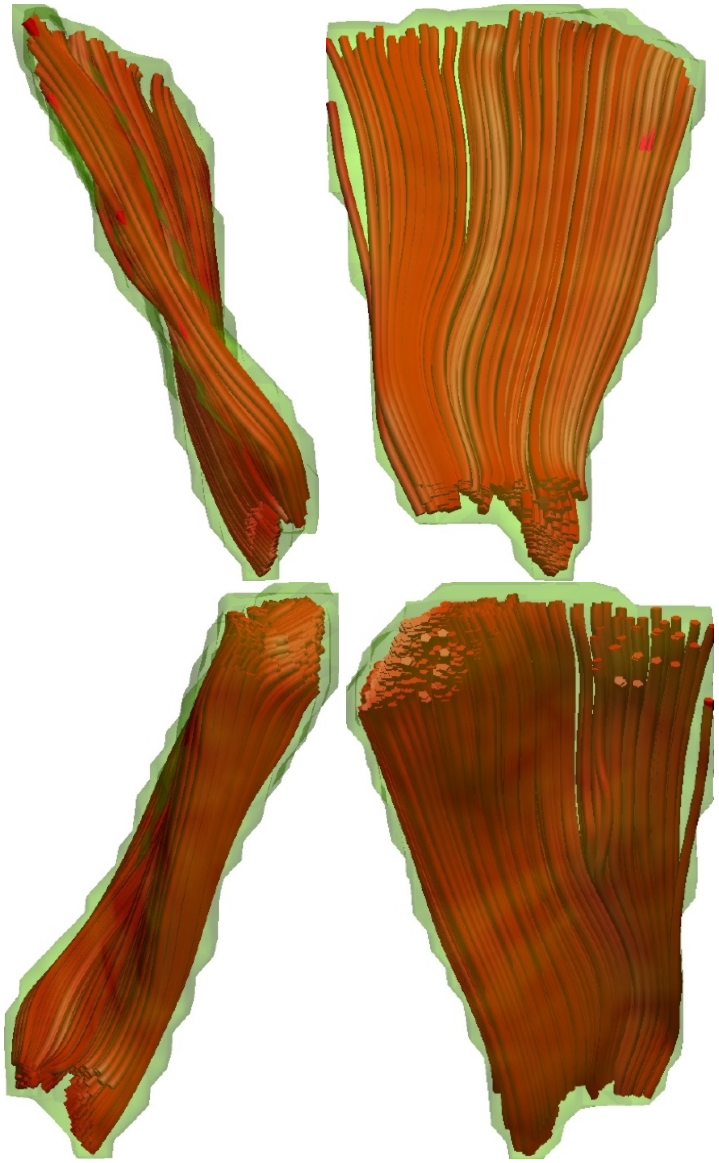

Figure 8.32: Four views of an SGM volume and modeling of the right corticospinal tract of an autism subject. Green shows voxels overlaying the SGM volume and SGM $\mathcal{C}_{\|}$ curves through the volume are shown in red.

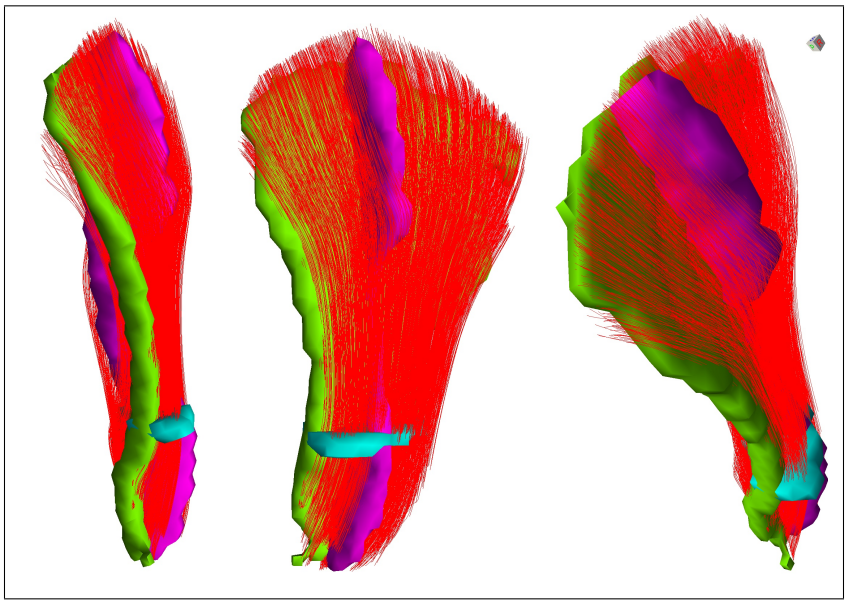

Figure 8.33: Corticospinal track: Three views of the right corticospinal tract with 3 orthogonal SGM surfaces and $\mathcal{C}_{||}$ curves.

8.10 Analyzing the Right Corticospinal tract – Sampling Method 1

The right corticospinal tract is one of the regions that has been shown to exhibit differences in white matter composition between people with normal development and people with Autism Spectrum Disorder [26]. Sampling Method 1, as depicted in Figure 8.10, was used to generate the functional data for this analysis. This resulted in sixteen smoothed max(FA) curves representing the control group and thirteen smoothed max(FA) curves representing the autism group. The curves were registered to a mean curve and the results plotted in Figure 8.34. T-tests compared the curves generated using Sampling Method 1.

No significant difference was found between the control and autism groups when simply comparing the distribution of max(FA) values along the curves. Additionally, the $F'(t)$ and $F''(t)$ curves showed no statistically significant differences at the .01 significance level on the right side of genu.

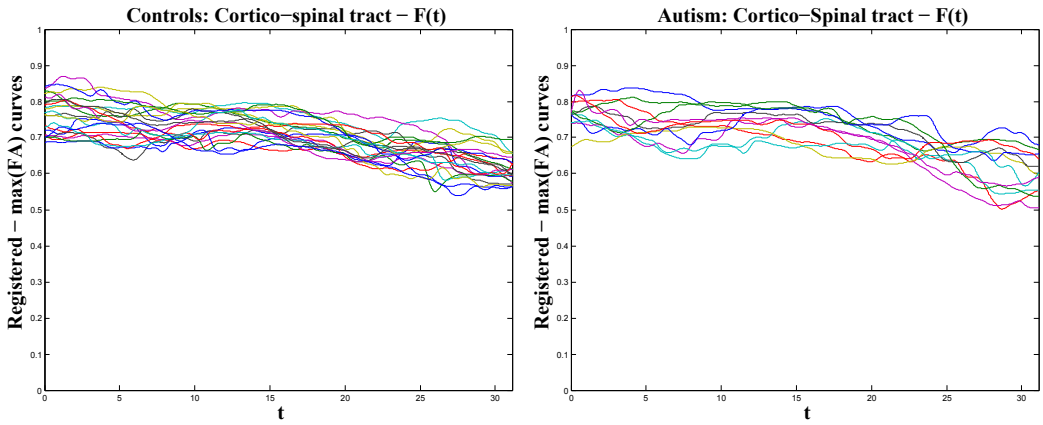

Figure 8.34: Corticospinal track: $F(t) = max(FA(t))$ curves for (left) control subjects, and (right) autism subjects.

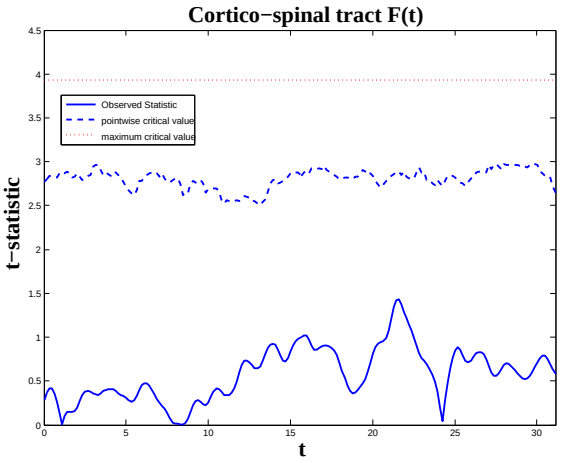

Figure 8.35: Corticospinal track: Max(FA) values Functional t-test controls vs autism for F(t). For the functional t-test there was no statistically significant difference between the control and autism curves.

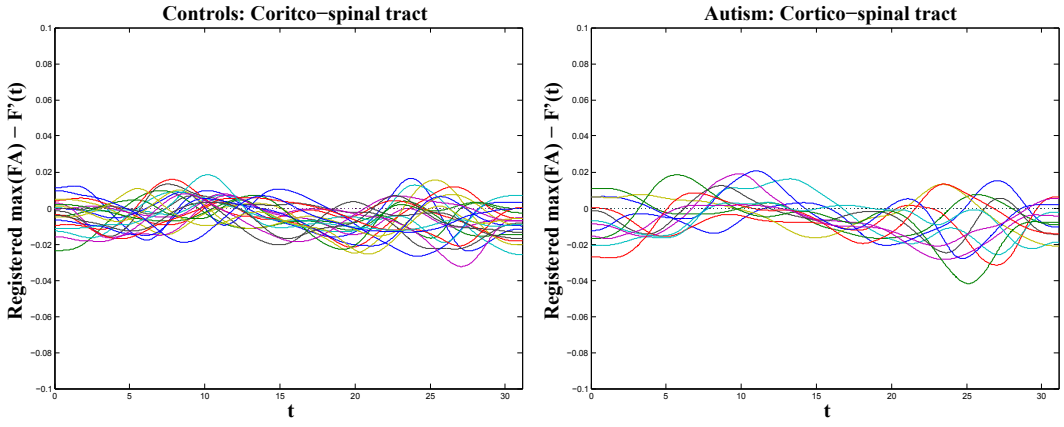

Figure 8.36: Corticospinal track: $F'(t) = \Delta max(FA(t))$ curves for (left) control subjects, and (right) autism subjects.

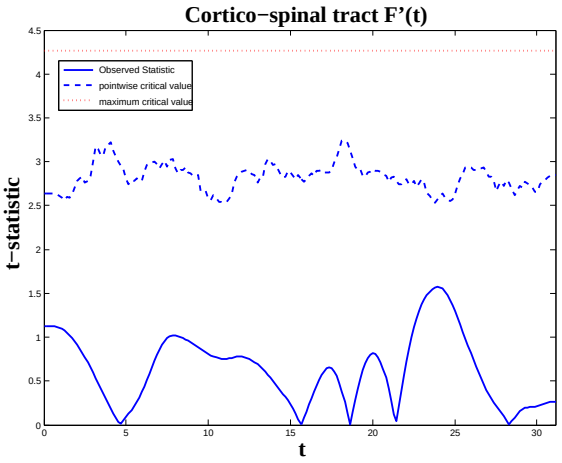

Figure 8.37: Corticospinal track $F'(t)$: Functional t-test for controls vs autism. For the functional t-test there was no statistically significant difference between the control and autism curves.

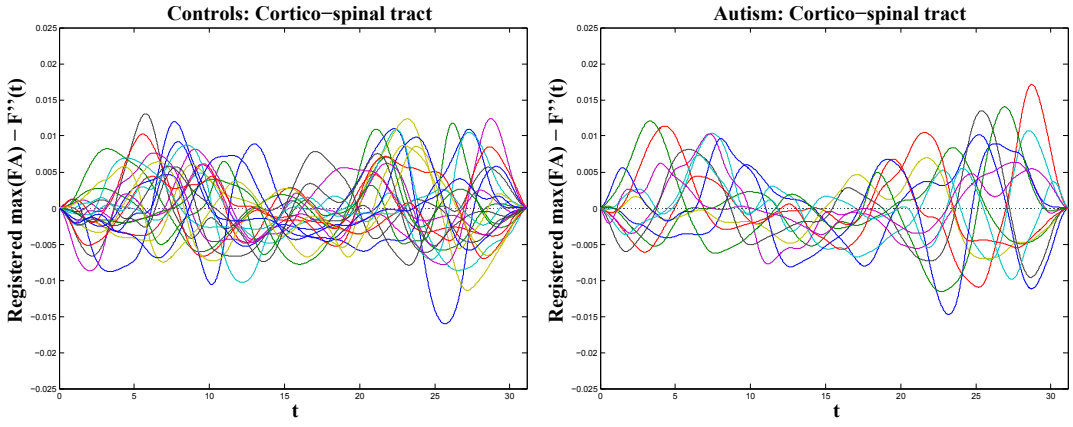

Figure 8.38: Corticospinal track: $F''(t) = \Delta^2 max(FA(t)$ curves for (left) control subjects, and (right) autism subjects.

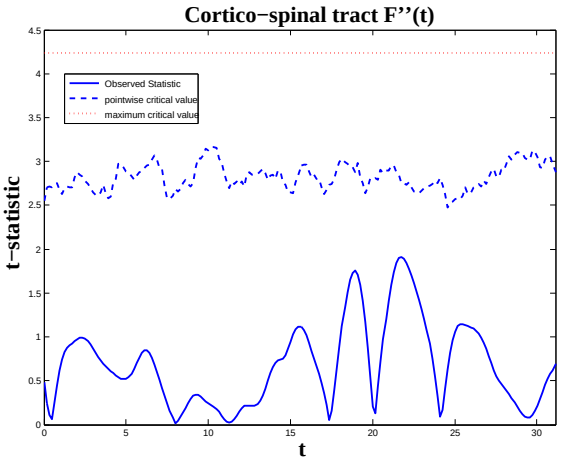

Figure 8.39: Corticospinal track $F''(t)$: Functional t-test Controls vs autism. For the functional t-test there were no statistically significant differences between the control and autism curves.

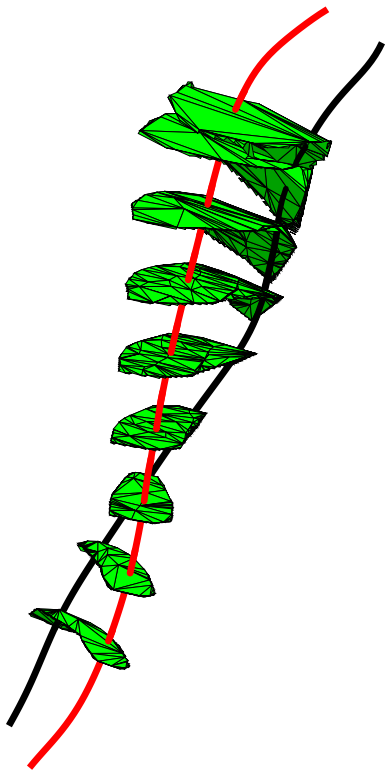

Figure 8.40: Corticospinal track: Sampling Method 2 Sampling surfaces were generated by: (1) taking t uniform steps along the $\mathcal{C}_{\parallel}^{M}(t)$ curves and (2) at each step t an orthogonal surface, $\mathcal{S}_{\parallel}(t)$ was calculated. Additional curves were randomly generated. FA values were mapped onto each curve based on the the locations determined by the intersection of that curve and the sampling surfaces.

8.11 Analyzing the Right Corticospinal Tract – Sampling Method 2

Sampling Method 2, as depicted in Figure 8.40 was used to generate the functional data curves, $F(t)$, for this analysis. The curves from the control group were bulked together and a random subset of 1000 curves was selected. The curves from the autism group were also bulked and a random subset of 1000 curves was selected. The curves were registered to the mean curve and the results plotted in Figure 8.20. T-tests were the used to test for statistically significant differences between the two groups. The $F'(t)$ and $F''(t)$ curves were generated by taking the first and second derivatives of the $F(t)$ curve respectively.

$F(t)$ showed significant differences between control and autism groups in the superior and inferior portions of the tract. $F'(t)$ showed statistically significant differences at the .01 significance in the superior, middle and inferior portions of the tract. $F''(t)$ also showed statistically significant differences at the .01 significance at three areas in the tract. The alignment of the volumes of the subjects as depicted in Figures 8.26 and 8.27 was not very good. It's also clear that the large ROIs used to define the tract contained more than one structure. The biological interpretation of the these differences remains to be determined.

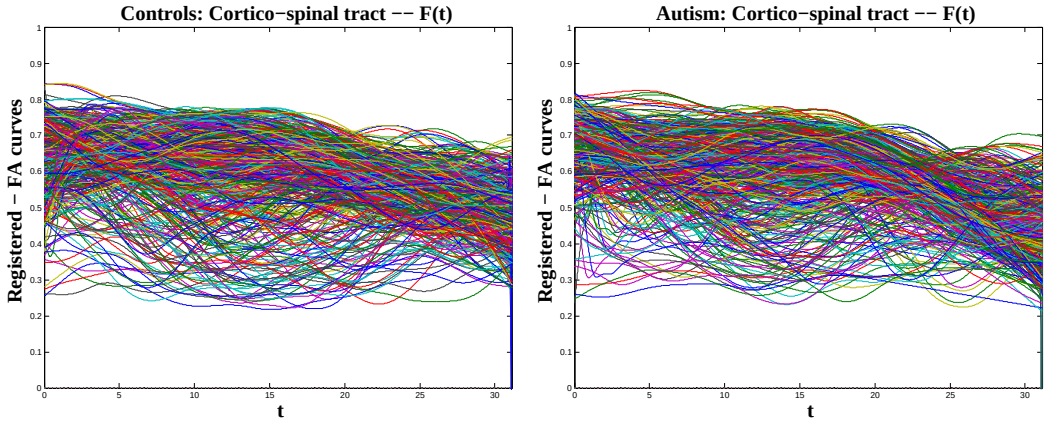

Figure 8.41: Corticospinal track: FA $F(t)$ curves for (left) control subjects, and (right) autism subjects.

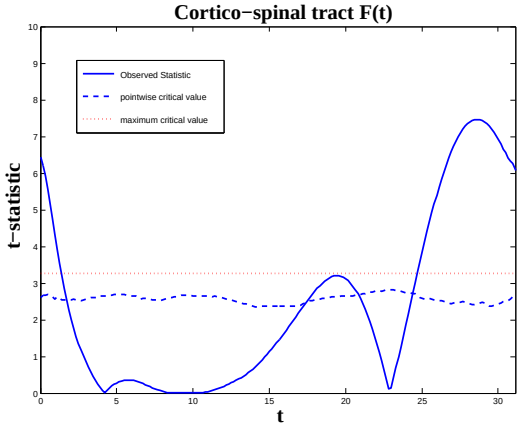

Figure 8.42: Corticospinal track full sampled volume: $F(t) = FA(t)$. Pairwise Functional t-test Controls vs autism for $F(t)$. The functional t-test finds statistically differences between the control and autism curves at the .01 significance level in the superior and inferior portions of the tract.

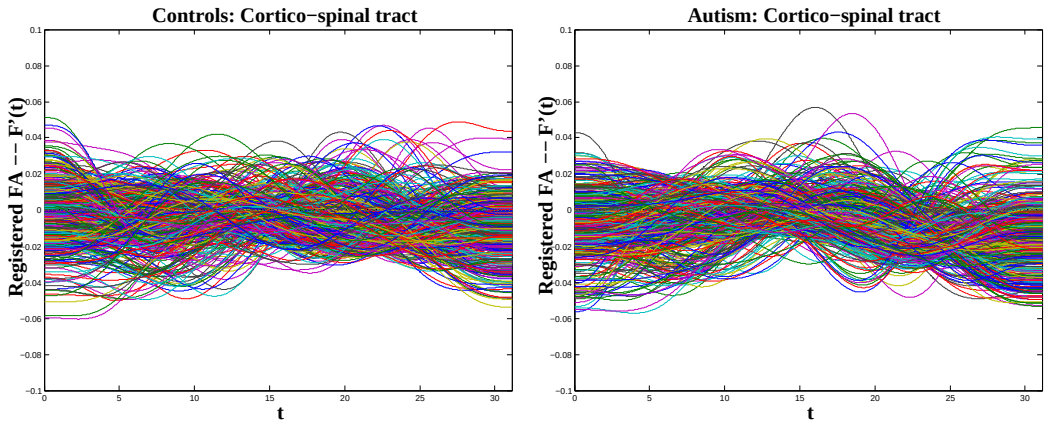

Figure 8.43: Corticospinal track: $F'(t) = \Delta FA(t)$ curves for (left) control subjects, and (right) autism subjects.

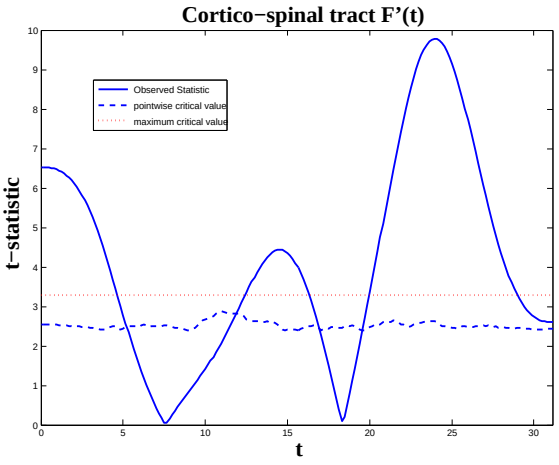

Figure 8.44: Corticospinal track full sampled volume: $F'(t)$ pairwise functional t-test controls vs autism groups. The t-test showed statistically significant differences at the .01 significance level in the superior, middle and inferior regions of the tract.

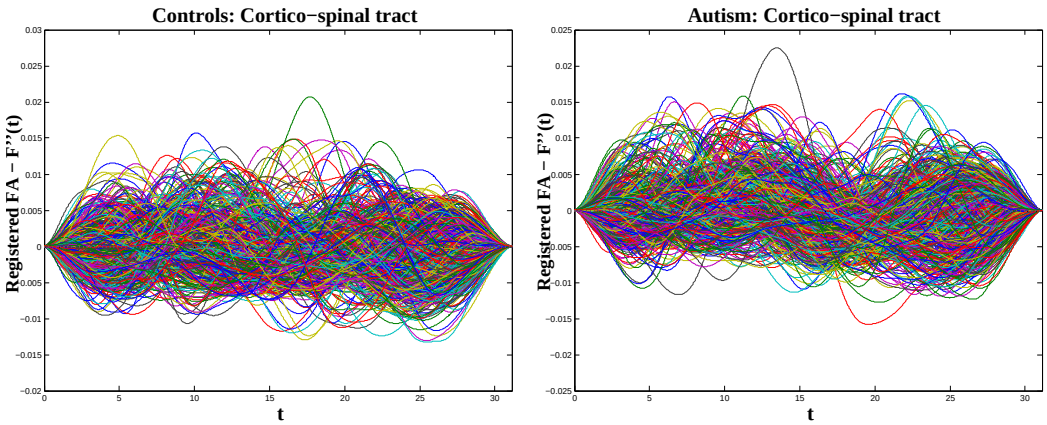

Figure 8.45: Corticospinal track full sampled volume: $F''(t) = \Delta^2 FA(t)$, curves for (left) control subjects, and (right) autism subjects.

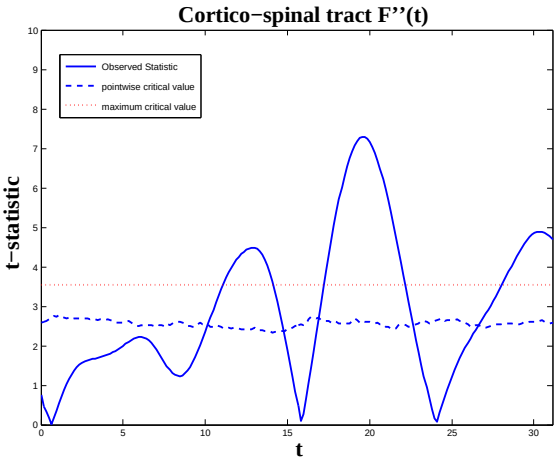

Figure 8.46: Corticospinal track full sampled volume: $F''(t)$ pairwise functional t-test for controls vs autism for $F''(t)$. The functional t-test finds statistically differences between the control and autism curves at the .01 significance level in three portions of the tract.

8.12 Discussion

This chapter presented SGM design choices that went into building of software suite an implementation of SGM. The software suite was used to analyze a Diffusion MRI data set and is able to detect, and localize group differences in two white matter structures that have been implicated in Autism Spectrum Disorder. Further work needs to be done to confirm that the analysis is biologically meaningful.

The suite included software to (1) build SGM models, (2) implements sampling and mapping functions, (3) outputs SGM volumes at voxel(in NIFTI format) and subvoxel levels (in Matlab mat format), (4) outputs SGM curves, $\mathcal{C}_{\parallel}, \mathcal{C}_{\perp}, \mathcal{C}_{\shortmid\!\vdash}$ in Camino, Trackvis formats, and Matlab mat format (5) maps users choice of invariants such as FA onto SGM curves and outputs each curve as a spline function ready for Functional Data Analysis, and (6) provides a wrapper around third party FDA package to do the analysis and generate reports.

Chapter 9

Discussion and Future Work

This thesis presented a new approach to modeling white matter structures Semiparametric Geometric Modeling. Semiparametric Geometric Modeling is a natural foundation for partitioning, modeling and analyzing the local, global, and volumetric properties of white matter structure. Unlike all current methods, the SGM inculcates the full structural information available from the data and from assumptions based on the global organization of white matter in the brain into a nonlinear manifold embedded in $\mathbb{R}^3$. This geometric model can be queried to extract geometric and scalar invariants based on the geometry of the white matter rather than on the imaging organization of the data's raw 3D rectilinear grid. The model can partition white matter structures into well defined geometric volumes, surfaces, curves, and points.

9.1 Review of Major Contributions

1. *A novel method for embedding a nonlinear manifold defined by a second order tensor field in* R^3.

2. *A novel method for modeling the foliations of a manifold defined by a vector field by a single implicit polynomial function that represents a family of implicit surfaces. This problem has previously been formulated as a*

local dynamic problem and solved by PDEs. That method suffers from sensitivity to noise and ignores global structure. Instead, the problem is formulated here as a static global problem that is solved by fitting the gradients of a function with global basis functions and "localizing" the function using the Geometric Median Surface. Chapters 4 and 5 laid out the basics of foliation theory as applied to vector fields. Modeling of foliations by families of implicit functions was introduced as well. Finally, the use Geometric Median Surface as a method of localizing the implicit functions to the data was laid out and justified.

3. *A novel robust depth function called a Geometric Median Surface (GMS). The GMS generalizes the geometric median point to parametric surfaces in vector fields.* Chapter 5 introduced the Geometric Median Surface as a depth function and introduced its use in localizing implicit polynomials to the manifold that the data was sampled from.

4. *A semi-supervised Support Vector method for clustering data on manifolds. This is an extension of the Slab SVM to a single class classifier. It clusters in a manner similar to single-link hierarchical clustering. Data is iteratively added to a labeled class based on compact kernel functions.* Chapters 5 and 6 defined the Semi-supervised Slab SVM and motivated the use of compact kernels the provided local similarity functions for clustering data from multiple manifolds to the specific manifold that they were sampled from. Chapter 8 presented specific kernels and evaluated the use of this method in an analysis of a Diffusion Weighted MRI data set.

5. *The SGM manifold model whose local structure is determined by a semi-supervised support vector model and whose global volumetric structure is a function of global, intersecting families of implicit surfaces. For example, in $\mathbb{R}^3$ for a manifold determined by a second-order tensor, the global volumetric structure amounts to the intersections of three families of orthogonally-intersecting implicit functions.* Chapters 5 described the

formulation of the SGM and and Chapter 6 presented the multi-objective optimization approach to fitting the SGM to the data. Chapter 8 presented the formulation of kernels, implicit algebraic polynomial, smoothing functions and constraint functions. The SGM was evaluated in as set of experiments that analyzed a Diffusion Weighted MRI data set.

6. *A structural query method to organize data on an SGM manifold for statistical analysis. For example Functional Data Analysis to analyze fractional anisotropy (FA) data along white matter fibers.* Chapter 7 introduced methods for querying geometric properties of the SGM. Additionally this chapter outlined the methods for mapping information from the rectilinear voxel space to the nonlinear manifold space of the white matter structures. Connectivity was

7. *A model that allows analytic calculation of local torsion, curvature and other differential geometric quantities at any point in the volume modeled by an SGM.* Chapter 7 introduced methods for querying geometric properties of the SGM

8. *A model that allows for rigid body transformations, for example rotations and translations, of an extracted volumetric structure by linear transforms of the coefficients of the implicit functions that are the output of the SGM.* Chapter 7 introduced methods for querying geometric properties of the SGM

9. *A new definition of white matter connectivity is defined based on the SGM's ability to partition white matter structures into well defined volumes.* Chapter 7 introduced methods for querying geometric properties of the SGM. Those properties include well defined volumes that can be used to partition white matter connecting specific region into non-overlapping volumes.

9.2 Future Work

The Diffusion Tensor has been a useful starting point for demonstrating SGMs it is only the simplest example. More sophisticated imaging techniques can provide more complete information about the local organization of white matter tissue that can be used to extend the SGM into regions of the brain where Diffusion Tensor information is ambiguous. Additionally, it would be useful to extend both the analysis techniques and modeling techniques that relax the orthogonality requires for local organization.

Also, recall that element of the internal structure of the manifold have a well defined total ordering. This is a valuable property when attempting to register volumes. Any registration method should not change to ordering of for example surfaces in the volume. If it did then the method has in some sense creased or torn the manifold. As part of future work a registration method based on smooth nonlinear warping of SGM manifolds will be explored.

Modeling manifolds defined by tensor fields is a very general approach. For example other venues for this approach are multidimensional wind velocity patterns, foliation patterns in metamorphic and sedimentary rock formations. Also this technique would be useful in geoimaging techniques where, like diffusion imaging the data can be very noisy but the is significant underlying structure.

Bibliography

[1] I. Aganj, C. Lenglet, N. Jahanshad, E. Yacoub, N. Harel, P. Thompson, and G. Sapiro, "A Hough transform global probabilistic approach to multiple-subject diffusion MRI tractography," in *Medical Image Analysis*, vol. 15, Elsevier, 2011, pp. 414–425.

[2] I. Aganj, C. Lenglet, R. Keriven, G. Sapiro, N. Harel, and P. Thompson, "A Hough transform global approach to diffusion MRI tractography," in *Proc. 17th International Society for Magnetic Resonance in Medicine*, vol. 17, 2009, p. 854.

[3] M. Aharon, M. Elad, and A. Bruckstein, "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[4] A. Aldroubi, C. Cabrelli, and U. Molter, "Optimal non-linear models for sparsity and sampling," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 793–812, 2008.

[5] A. Alexander, J. Lee, M. Lazar, R. Boudos, M. DuBray, T. Oakes, J. Miller, J. Lu, E. Jeong, W. McMahon, E. Bigler, and J. Lainhart, "Diffusion tensor imaging of the corpus callosum in autism," *Neuroimage*, vol. 34, no. 1, pp. 61–73, 2007.

[6] D. Alexander and G. Barker, "Optimal imaging parameters for fibre-orientation estimation in Diffusion MRI," *NeuroImage*, vol. 27, pp. 357–367, 2005.

[7] G. Aloupis, "Geometric measures of data depth," *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 72, p. 147, 2006.

[8] S. Amari, "Geometry of semiparametric models and applications," *International Statistical Institute*,

[9] M. Arnaudon and F. Nielsen, "Medians and means in Finsler geometry," *Arxiv preprint arXiv:1011.6076*, 2010.

[10] J. Ashburner and K. Friston, "Voxel-based morphometry—-The methods," *Neuroimage*, vol. 11, no. 6, pp. 805–821, 2000.

[11] C. Bajaj, *Introduction to Implicit Surfaces*. Morgan Kaufmann, 1997.

[12] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "In vivo fiber tractography using DT-MRI data," *Magnetic Resonance in Medicine*, vol. 44, no. 4, pp. 625–632, 2000.

[13] P. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *Journal of Magnetic Resonance-Series B*, vol. 111, no. 3, pp. 209–219, 1996.

[14] T. Behrens, M. Woolrich, M. Jenkinson, H. Johansen-Berg, R. Nunes, S. Clare, P. Matthews, J. Brady, and S. Smith, "Characterization and propagation of uncertainty in Diffusion-Weighted MR imaging," *Magnetic Resonance in Medicine*, vol. 50, no. 5, pp. 1077–1088, 2003.

[15] M. Belkin, "Problems of Learning on Manifolds," PhD thesis, Department of Mathematics, University of Chicago, 2003.

[16] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[17] O. Ben-Shahar and S. Zucker, "The perceptual organization of texture flow: A contextual inference approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 401–417, 2003.

[18] H. Benveniste, L. Hedlund, and G. Johnson, "Mechanism of detection of acute cerebral ischemia in rats by diffusion-weighted magnetic resonance microscopy," *Stroke*, vol. 23, no. 5, pp. 746–754, 1992.

[19] O. Bergmann, A. Lundervold, and T. Steihaug, "Generating a synthetic Diffusion Tensor dataset," *Proc. 18th IEEE Symposium on Computer-Based Medical Systems*, pp. 277–281, 2005.

[20] P. Bickel and B. Li, "Local polynomial regression on unknown manifolds," *Complex Datasets and Inverse Problems: Tomography, Networks and Beyond. Institute of Mathematical Statistics Lecture Notes-Monograph Series*, vol. 54, pp. 177–186, 2007.

[21] R. Bishop and S. Goldberg, "Tensor Analysis on Manifolds," *Dover Publication, NY*, 1980.

[22] M. Blane, Z. Lei, H. Civi, and D. Cooper, "Algorithm for fitting implicit polynomial curves and surfaces to data," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 22, no. 3, pp. 298–313, 2000.

[23] H. Blum and R. Nagel, "Shape description using weighted symmetric axis features," *Pattern Recognition*, vol. 10, no. 3, pp. 167–180, 1978.

[24] N. Bose and N. Ahuja, "Superresolution and noise filtering using moving least squares," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2239–2248, 2006.

[25] P. Bradley and O. Mangasarian, "k-plane clustering," *Journal of Global Optimization*, vol. 16, no. 1, pp. 23–32, 2000.

[26] A. Brito, M. Vasconcelos, R. Domingues, L. Hygino da Cruz Jr, L. Rodrigues, E. Gasparetto, and C. Calçada, "Diffusion tensor imaging findings in school-aged autistic children," *Journal of Neuroimaging*, vol. 19, no. 4, pp. 337–343, 2009.

[27] C. Burges, "A tutorial on Support Vector Machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[28] B. Cabral and L. Leedom, "Imaging vector fields using line integral convolution," in *Proc. 20th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 1993, pp. 263–270.

[29] C. Camacho and A. Lins Neto, *Geometric Theory of Foliations*. Boston: Birkhäuser, 1985.

[30] J. Campbell, K. Siddiqi, V. Rymar, A. Sadikot, and G. Pike, "Flow-based fiber tracking with diffusion tensor and q-ball data: validation and comparison to principal diffusion direction techniques," *NeuroImage*, vol. 27, no. 4, pp. 725–736, 2005.

[31] O. Chapelle and B. Schölkopf, "Incorporating invariances in non-linear Support Vector Machines," *Advances in Neural Information Processing Systems*, vol. 1, pp. 609–616, 2002.

[32] O. Chapelle and A. Zien, "Semi-supervised classification by low density separation," in *Proc. of the Tenth International Workshop on Artificial Intelligence and Statistics*, Society for Artificial Intelligence and Statistics, 2005, pp. 57–64.

[33] G. Chirikjian and O. Chétalet, "Sampling and convolution on motion groups using generalized Gaussian functions," *Electronic Journal of Computational Kinematics*, vol. 1, no. 1, 2002.

[34] M. Chung, *Computational Neuroanatomy: The Methods*. World Scientific, 2012.

[35] M. Chung, J. Lee, G. Pack, M. Lazar, N. Lange, J. Lainhart, and L. Alexander, "A unified parametric model of white matter fiber tracts," *MACCAI Workshop on Computational Diffusion MRI*, 2008.

[36] M. Chung, N. Adluru, J. Lee, M. Lazar, J. Lainhart, and A. Alexander, "Efficient parametric encoding scheme for white matter fiber bundles," in *Proc. IEEE International Conference Engineering in Medicine and Biology Society,*, 2009, pp. 6644–6647.

[37] A. Clauset, C. Moore, and M. Newman, "Structural inference of hierarchies in networks," in *Statistical Network Analysis: Models, Issues, and New Directions*, Springer, 2007, pp. 1–13.

[38] E. Coddington, N. Levinson, and T. Teichmann, "Theory of ordinary differential equations," *Physics Today*, vol. 9, p. 18, 1956.

[39] P. A. Cook, Y. Bai, N. S. Gilani, K. K. Seunarine, M. G. Hall, G. J. Parker, and D. C. Alexander, "Camino: open-source diffusion-MRI reconstruction and processing," in *14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, vol. 2759, 2006.

[40] O. Coulon, D. Alexander, and S. Arridge, "A regularization scheme for Diffusion Tensor Magnetic Resonance Images," *Proc. Information Processing in Medical Imaging*, pp. 92–105, 2001.

[41] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass SVMs," *Journal of Machine Learning Research*, vol. 2, pp. 265–292, 2001.

[42] P. Craven and G. Wahba, "Smoothing noisy data with spline functions," *Numerische Mathematik*, vol. 31, no. 4, pp. 377–403, 1978.

[43] K. Dalton, B. Nacewicz, T. Johnstone, H. Schaefer, M. Gernsbacher, H. Goldsmith, A. Alexander, and R. Davidson, "Gaze fixation and the neural circuitry of face processing in autism," *Nature Neuroscience*, vol. 8, no. 4, pp. 519–526, 2005.

[44] C. Demetrescu, A. Goldberg, and D. Johnson, *The shortest path problem: Ninth DIMACS implementation challenge*. American Mathematical Society, 2009, vol. 74, pp. 41–72.

[45] A. Desolneux, L. Moisan, and J. Morel, "Meaningful alignments," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 7–23, 2000.

[46] L. Devroye, *Non-Uniform Random Variate Generation*. Springer-Verlag New York, 1986, vol. 4.

[47] E. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[48] P. Dollár, V. Rabaud, and S. Belongie, "Learning to traverse image manifolds," in *Proc. Advances in Neural Information Processing Systems*, vol. 19, 2007, p. 361.

[49] R. F. Dougherty, M. Ben-Shachar, R. Bammer, A. A. Brewer, and B. A. Wandell, "Functional organization of human occipital-callosal fiber tracts," *Proc. National Academy of Sciences*, vol. 102, no. 20, pp. 7350–7355, 2005.

[50] M. Eigensatz, J. Giesen, and M. Manjunath, "The solution path of the Slab Support Vector Machine," in *Proc. 20th Canadian Conference on Computational Geometry, McGill University*, 2008, pp. 211–214.

[51] L. Eisenhart, *A Treatise on the Differential Geometry of Curves and Surfaces*. Ginn and Company, 1909.

[52] H. Eyal-Giladi, "Establishment of the axis in chordates: Facts and speculations," *Development*, vol. 124, no. 12, pp. 2285–2296, 1997.

[53] N. Fisher, T. Lewis, and B. Embleton, *Statistical Analysis of Spherical Data*. Cambridge University Press, 1993.

[54] P. Fletcher, S. Joshi, C. Lu, and S. Pizer, "Gaussian distributions on Lie groups and their application to statistical shape analysis," in *Information Processing in Medical Imaging*, Springer, 2003, pp. 450–462.

[55] P. Fletcher, R. Tao, W. Jeong, and R. Whitaker, "A volumetric approach to quantifying region-to-region white matter connectivity in diffusion tensor mri," in *Information Processing in Medical Imaging*, Springer, 2007, pp. 346–358.

[56] P. Fletcher and S. Joshi, "Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors," in *Lecture Notes in Computer Science*, Springer, 2004, pp. 87–98.

[57] ——, "Riemannian geometry for the statistical analysis of Diffusion Tensor data," *Signal Processing*, vol. 87, no. 2, pp. 250–262, 2007.

[58] D. Freedman, *Statistical Models: Theory and Practice*. Cambridge University Press, 2005.

[59] K. Gauvain, R. McKinstry, P. Mukherjee, A. Perry, J. Neil, B. Kaufman, and R. Hayashi, "Evaluating pediatric brain tumor cellularity with diffusion-tensor imaging," *American Journal of Roentgenology*, vol. 177, no. 2, pp. 449–454, 2001.

[60] G. Gerig, I. Corouge, C. Vachet, K. Krishnan, and J. MacFall, "Quantitative analysis of diffusion properties of white matter fiber tracts: A validation study," *13th Proc. of the International Society for Magnetic Resonance in Medicine*, p. 1337, 2005.

[61] A. B. Goldberg, X. Zhu, A. Singh, Z. Xu, and R. Nowak, "Multi-manifold semi-supervised learning," in *Proc. 12th Int. Conf. Artificial Intell. and Statistics*, 2009, 1074–1086.

[62] Y. Goldberg, A. Zakai, D. Kushnir, and Y. Ritov, "Manifold learning: The price of normalization," *J. Machine Learning Research*, vol. 9, pp. 1909–1939, 2008.

[63] R. Goldman, "Curvature formulas for implicit curves and surfaces," *Computer Aided Geometric Design*, vol. 22, no. 7, pp. 632–658, 2005.

[64] J. Goldsmith, B. Caffo, C. Crainiceanu, D. Reich, Y. Du, and C. Hendrix, "Nonlinear tube-fitting for the analysis of anatomical and functional structures," *The Annals of Applied Statistics*, vol. 5, no. 1, p. 337, 2011.

[65] J. Goldsmith, C. Crainiceanu, B. Caffo, and D. Reich, "Penalized functional regression analysis of white-matter tract profiles in multiple sclerosis," *NeuroImage*, vol. 57, no. 2, pp. 431–439, 2011.

[66] J. Goldsmith, J. Feder, C. Crainiceanu, B. Caffo, and D. Reich, "Penalized functional regression," *Johns Hopkins University, Dept. of Biostatistics Working Papers*, p. 204, 2010.

[67] G. Gong, T. Jiang, C. Zhu, Y. Zang, F. Wang, S. Xie, J. Xiao, and X. Guo, "Asymmetry analysis of cingulum based on scale-invariant parameterization by diffusion tensor imaging," *Human Brain Mapping*, vol. 24, no. 2, pp. 92–98, 2004.

[68] C. Goodlett, P. Fletcher, J. Gilmore, and G. Gerig, "Group analysis of DTI fiber tract statistics with application to neurodevelopment," *NeuroImage*, vol. 45, no. 1, S133–S142, 2009.

[69] N. Hageman, A. Toga, K. Narr, and D. Shattuck, "A diffusion tensor imaging tractography algorithm based on Navier–Stokes fluid mechanics," *IEEE Transactions on Medical Imaging*, vol. 28, no. 3, pp. 348–360, 2009.

[70] P. Hagmann, M. Kurant, X. Gigandet, P. Thiran, V. J. Wedeen, R. Meuli, and J. P. Thiran, "Mapping human whole-brain structural networks with diffusion MRI," *PLoS ONE*, vol. 2, no. 7, p. 597, 2007.

[71] A. Hanson, "Quaternion Frenet frames: Making optimal tubes and ribbons from curves," Indiana University Computer Science Department, Tech. Rep. 407, 1994.

[72] G. Haro, C. Lenglet, G. Sapiro, and P. Thompson, "On the non-uniform complexity of brain connectivity," *Proc. 5th IEEE International Symposium on Biomedical Imaging From Nano to Macro*, pp. 887–890, 2008.

[73] T. Hastie and W. Stuetzle, "Principal curves," *Journal of the American Statistical Association*, vol. 84, no. 406, pp. 502–516, 1989.

[74] G. Hewer, C. Kenney, and B. Manjunath, "Image segmentation via functionals based on boundary functions," in *Proc. IEEE International Conference on Image Processing*, vol. 1, 1996, pp. 813–816.

[75] D. Hochbaum, "An efficient algorithm for image segmentation, Markov Random Fields and related problems," *Journal of the ACM*, vol. 48, no. 4, pp. 686–701, 2001.

[76] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, "Surface reconstruction from unorganized points," *Proc. of ACM SIGGRAPH*, vol. 26, pp. 71–78, 1992.

[77] R. Hummel and S. Zucker, "On the foundations of relaxation labeling processes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 3, pp. 267–287, 1983.

[78] M. Irani and S. Peleg, "Super resolution from image sequences," *Proc. International Conference on Pattern Recognition-C*, vol. 90, pp. 115–120, 1990.

[79] S. Jbabdi, M. W. Woolrich, J. L. R. Andersson, and T. E. J. Behrens, "A Bayesian framework for global tractography," *Neuroimage*, vol. 37, no. 1, pp. 116–129, 2007.

[80] T. Joachims, "Making large-scale support vector machine learning practical," in *Advances in Kernel Methods*, MIT Press, 1999, pp. 169–184.

[81] H. Johansen-Berg and T. Behrens, *Diffusion MRI: From quantitative measurement to in-vivo neuroanatomy*. Academic Press, 2009, ch. 7, pp. 127–146.

[82] ——, *Diffusion MRI: from quantitative measurement to in-vivo neuroanatomy*. Academic Press, 2009, ch. 15, p. 348.

[83] L. Jonasson, P. Hagmann, J. Thiran, and V. Wedeen, "Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering," in *International Society for Magnetic Resonance in Medicine*, vol. 24, 2005, pp. 1127–1137.

[84] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, no. 1, p. 151, 2004.

[85] N. Kang, J. Zhang, E. Carlson, and D. Gembris, "White matter fiber tractography via anisotropic diffusion simulation in the human brain," *IEEE Transactions on Medical Imaging*, vol. 24, no. 9, pp. 1127–1137, 2005.

[86] B. Kimia, I. Frankel, and A. Popescu, "Euler spiral for shape completion," *International Journal of Computer Vision*, vol. 54, no. 1, pp. 159–182, 2003.

[87] R. Kolluri, "Provably good moving least squares," *ACM Transactions on Algorithms*, vol. 4, no. 2, p. 18, 2008.

[88] S. König, S. Gumhold, and T. Dresden, "Consistent propagation of normal orientations in point clouds," *Proc. Vision, Modeling, and Visualization Workshop*, 2009.

[89] B. W. Kreher, I. Mader, and V. G. Kiselev, "Gibbs tracking: A novel approach for the reconstruction of neuronal pathways," *Magnetic Resonance in Medicine*, vol. 60, pp. 953–963, 2008.

[90] B. Kreher, J. Schneider, I. Mader, E. Martin, J. Hennig, and K. Il'yasov, "Multitensor approach for analysis and tracking of complex fiber configurations," *Magnetic Resonance in Medicine*, vol. 54, pp. 1216–1225, 2005.

[91] A. Kyatkin and G. Chirikjian, "Synthesis of binary manipulators using the Fourier Transform on the Euclidean Group," *Journal of Mechanical Design*, vol. 121, p. 9, 1999.

[92] F. Lauer and G. Bloch, "Incorporating prior knowledge in Support Vector Machines for classification: A review," *Neurocomputing*, vol. 71, no. 7-9, pp. 1578–1594, 2008.

[93] M. Lazar and A. L. Alexander, "Bootstrap white matter tractography (BOOT-TRAC)," *Neuroimage*, vol. 24, no. 2, pp. 524–532, 2005.

[94] M. Lazar and A. Alexander, "An error analysis of white matter tractography methods: Synthetic diffusion tensor field simulations," *Neuroimage*, vol. 20, no. 2, pp. 1140–1153, 2003.

[95] M. Lazar, J. Lee, and A. Alexander, "Axial asymmetry of water diffusion in brain white matter," *Magnetic Resonance in Medicine*, vol. 54, no. 4, pp. 860–867, 2005.

[96] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. E. Meyerand, B. Badie, H. A. Rowley, V. Haughton, A. Field, and A. L. Alexander, "White matter tractography using diffusion tensor deflection," *Human Brain Mapping*, vol. 18, no. 4, pp. 306–321, 2003.

[97] J. Lee and M. Verleysen, *Nonlinear Dimensionality Reduction*. New York: Springer, 2007.

[98] J. Lee, E. Bigler, A. Alexander, M. Lazar, M. DuBray, M. Chung, M. Johnson, J. Morgan, J. Miller, W. McMahon, J. Lu, E. Jeong, and J. Lainhart, "Diffusion tensor imaging of white matter in the superior temporal gyrus and temporal stem in autism," *Neuroscience Letters*, vol. 424, no. 2, pp. 127–132, 2007.

[99] S. Lee and S. Wright, "Decomposition algorithms for training large-scale Semiparametric Support Vector Machines," in *Proc. of the European Conference on Machine Learning and Knowledge Discovery in Databases: Part II*, 2009, p. 14.

[100] Y. Lee, Y. Kim, S. Lee, and J. Koo, "Structured multicategory Support Vector Machine with analysis of variance decomposition," *Biometrika*, vol. 93, no. 3, pp. 555–571, 2006.

[101] H. Liu, R. Chellappak, and A. Rosenfeld, "Accurate dense optical flow estimation using adaptive structure tensors and parametric model," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1170–1180, 2003.

[102] S. López-Pintado and J. Romo, "On the concept of depth for functional data," *Journal of the American Statistical Association*, vol. 104, no. 486, pp. 718–734, 2009.

[103] C. Lord, M. Rutter, and A. Couteur, "Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders," *Journal of Autism and Developmental Disorders*, vol. 24, no. 5, pp. 659–685, 1994.

[104] R. Lordo, "Nonparametric and semiparametric models," *Technometrics*, vol. 47, no. 2, pp. 234–234, 2005.

[105] M. Maddah, W. Grimson, S. Warfield, and W. Wells, "A unified framework for clustering and quantitative analysis of white matter fiber tracts," *Medical Image Analysis*, vol. 12, no. 2, pp. 191–202, 2008.

[106] O. Mangasarian, J. Shavlik, and E. Wild, "Knowledge-based kernel approximation," *Journal of Machine Learning Research*, vol. 5, pp. 1127–1141, 2004.

[107] O. Mangasarian and E. Wild, "Nonlinear knowledge in kernel approximation," *IEEE Transactions on Neural Networks*, vol. 18, no. 1, p. 300, 2007.

[108] O. L. Mangasarian and E. W. Wild, "Multisurface proximal Support Vector Machine classification via generalized eigenvalues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 69–74, 2006.

[109] K. Mardia and P. Jupp, *Directional Statistics*. Wiley, 2009, vol. 494.

[110] MATLAB, *version 7.13.0.564 (R2011b)*. Natick, Massachusetts: The MathWorks Inc., 2011.

[111] G. Mayer and E. Vrscay, "Measuring information gain for frequency-encoded super-resolution MRI," *Magnetic Resonance Imaging*, vol. 25, no. 7, pp. 1058–1069, 2007.

[112] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis," in *Proc. of the Workshop on Mathematical Methods in Biomedical Image Analysis*, 1996, pp. 171–180.

[113] F. Memoli and G. Sapiro, "Distance functions and geodesics on submanifolds of $\mathbb{r}^{\mathrm{d}}$ and point clouds," *SIAM Journal on Applied Mathematics*, vol. 65, no. 4, pp. 1227–1260, 2005.

[114] K. Miettinen, *Nonlinear Multiobjective Optimization*. Springer, 1999, vol. 12.

[115] J. Milnor, "Foliations and Foliated Vector Bundles, Collected Papers of John Milnor: IV. Homotopy, Homology and Manifolds, ed. J. McCleary," *American Mathematical Society*, pp. 279–320, 2009.

[116] I. Moerdijk and J. Mrcun, *Introduction to Foliations and Lie Groupoids*. Cambridge, UK: Cambridge University Press, 2003.

[117] R. Mohan, R. Nevatia, I. Center, and Y. Heights, "Perceptual organization for scene segmentation and description," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 6, pp. 616–635, 1992.

[118] P. Mordohai and G. Medioni, "Dimensionality estimation, manifold learning and function approximation using tensor voting," *J. Machine Learning Research*, vol. 11, pp. 411–450, 2010.

[119] S. Mori and P. van Zijl, "Fiber tracking: Principles and strategies–a technical review," *NMR in Biomedicine*, vol. 15, no. 7-8, pp. 468–480, 2002.

[120] A. Narmanov and G. Kaypnazarova, "Foliation theory and its applications," *Turkic World Mathematical Society Journal of Pure and Applied Mathematics*, vol. 2, no. 1, pp. 112–126, 2011.

[121] C. V. Nguyen and D. B. H. Tay, "Regression using multikernel and Semiparametric Support Vector algorithms," *IEEE Signal Processing Letters*, vol. 15, pp. 481–484, 2008.

[122] N. Nguyen and P. Milanfar, "A wavelet-based interpolation-restoration method for superresolution (wavelet superresolution)," *IEEE Transactions on Circuits, Systems, and Signal Processing*, vol. 19, no. 4, pp. 321–338, 2000.

[123] T. Nishimori, "Behaviour of leaves of codimension-one foliations," *Tohoku Mathematical Journal*, vol. 29, no. 2, pp. 255–273, 1977.

[124] L. O'Donnell, S. Haker, and C.-F. Westin, "New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and geodesics in a tensor-warped space," *Proc. of 5th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 459–466, 2002.

[125] L. O'Donnell and C. Westin, "White matter tract clustering and correspondence in populations," *Proc. from Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 140–147, 2005.

[126] S. Osher and N. Paragios, *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer-Verlag, New York, 2003.

[127] G. Parker and D. Alexander, "Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information," in *Proc. Information Processing in Medical Imaging*, Springer, 2003, pp. 684–695.

[128] S. Peled and Y. Yeshurun, "Superresolution in MRI: Application to human white matter fiber tract visualization by Diffusion Tensor Imaging," *Magnetic Resonance in Medicine*, vol. 45, no. 1, pp. 29–35, 2001.

[129] ——, "Superresolution in MRI perhaps sometimes," *Magnetic Resonance in Medicine*, vol. 48, no. 2, pp. 409–409, 2002.

[130] S. Pizer, P. Fletcher, S. Joshi, A. Thall, J. Chen, Y. Fridman, D. Fritsch, A. Gash, J. Glotzer, and M. Jiroutek, "Deformable m-reps for 3D medical image segmentation," *International Journal of Computer Vision*, vol. 55, no. 2, pp. 85–106, 2003.

[131] J. Platt, "Fast training of Support Vector Machines using sequential minimal optimization," in *Advances in Kernel Methods*, MIT Press, 1999, pp. 185–208.

[132] A. Ramirez-Manzanares, M. Rivera, B. Vemuri, P. Carney, and T. Mareci, "Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry," *IEEE Trans. on Medical Imaging*, vol. 26, no. 8, pp. 1091–1102, 2007.

[133] J. Ramsay, *Functional Data Analysis*. Wiley Online Library, 2005.

[134] J. O. Ramsay, S. Graves, and G. Hooker, *Functional Data Analysis with R and MATLAB*. Springer, 2009.

[135] J. O. Ramsay and B. W. Silverman, *Applied Functional Data Analysis: Methods and Case Studies*. New York: Springer-Verlag, 2002.

[136] ——, *Functional Data Analysis*. Springer, 1997.

[137] J. Ramsay and X. Li, "Curve registration," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 60, no. 2, pp. 351–363, 2002.

[138] D. Rolfsen, "Knots and links, corrected reprint of the 1976 original," *Mathematics Lecture Series*, vol. 7, 1990.

[139] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[140] P. D. Sampson, "Fitting conic sections to very scattered data: An iterative refinement of the Bookstein algorithm," *Computer Graphics and Image Processing*, vol. 18, no. 1, pp. 97–108, 1982.

[141] L. K. Saul and S. T. Roweis, "Think globally, fit locally: Unsupervised learning of low dimensional manifolds," *Journal of Machine Learning Research*, vol. 4, pp. 119–155, 2003.

[142] K. Scheffler, "Superresolution in MRI?," *Magnetic Resonance in Medicine*, vol. 48, no. 2, pp. 408–408, 2002.

[143] B. Schöelkopf, P. Simard, A. Smola, and V. Vapnik, "Prior knowledge in support vector kernels," *Advances in Neural Information Processing Systems*, pp. 640–646, 1998.

[144] B. Schölkopf, J. Giesen, and S. Spalinger, "Kernel methods for implicit surface modeling," in *Advances in Neural Information Processing Systems 17*, L. Saul, Y. Weiss, and L. Bottou, Eds., Cambridge, MA: MIT Press, 2005, pp. 1193–1200.

[145] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[146] B. Schölkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, 2001.

[147] M. Searcóid, *Metric Spaces*. Springer Verlag, 2006.

[148] D. Shen, S. Moffat, S. Resnick, and C. Davatzikos, "Measuring size and shape of the hippocampus in MR images using a deformable shape model," *Neuroimage*, vol. 15, no. 2, pp. 422–434, 2002.

[149] S. Shields, "The simplest branched surfaces for a foliation," *Pacific Journal of Mathematics*, vol. 225, no. 1, p. 177, 2006.

[150] S. Smith, P. Bannister, C. Beckmann, M. Brady, S. Clare, D. Flitney, P. Hansen, M. Jenkinson, D. Leibovici, and B. Ripley, "FSL: New tools for functional and structural brain image analysis," *NeuroImage*, vol. 13, no. 6, pp. 249–249, 2001.

[151] S. Smith, M. Jenkinson, H. Johansen-Berg, D. Rueckert, T. Nichols, C. Mackay, K. Watkins, O. Ciccarelli, M. Cader, and P. Matthews, "Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data," *Neuroimage*, vol. 31, no. 4, pp. 1487–1505, 2006.

[152] A. J. Smola, T. T. Friess, and B. Schölkopf, "Semiparametric support vector and linear programming machines," in *Advances in Neural Information Processing Systems 11*, M. S. Kearns, S. A. Solla, and D. A. Cohn, Eds., Cambridge, MA: MIT Press, 1999, pp. 585–591.

[153] J. Sotomayor and R. Garcia, "Lines of curvature on surfaces, historical comments and recent developments," *Arxiv preprint arXiv:0712.1585*, 2007.

[154] A. Sparavigna and B. Montrucchio, "Vector field visualization with streamlines," *Arxiv preprint cs/0610088 Graphics*, pp. 1–9, 2006.

[155] M. Spivak and M. Spivak, *A Comprehensive Introduction to Differential Geometry*. Publish or Perish, Berkeley, 1979, vol. 1.

[156] O. Sporns, G. Tononi, and R. Kötter, "The human connectome: A structural description of the human brain," *PLoS Computational Biology*, vol. 1, no. 4, pp. 245–251, 2005.

[157] F. Steinke and B. Schölkopf, "Kernels, regularization and differential equations," *Pattern Recognition*, vol. 41, no. 11, pp. 3271–3286, 2008.

[158] M. Styner and G. Gerig, "Medial models incorporating object variability for 3D shape analysis," in *Proc. Information Processing in Medical Imaging*, Springer, 2001, pp. 502–516.

[159] S. Sun, H. Liang, K. Trinkaus, A. Cross, R. Armstrong, and S. Song, "Noninvasive detection of cuprizone induced axonal damage and demyelination in the mouse corpus callosum," *Magnetic Resonance in Medicine*, vol. 55, no. 2, pp. 302–308, 2006.

[160] H. Takeda, "Kernel regression for image processing and reconstruction," PhD thesis, University of California, 2006.

[161] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349–366, 2007.

[162] ——, "Robust kernel regression for restoration and reconstruction of images from sparse noisy data," *IEEE International Conference on Image Processing*, pp. 1257–1260, 2006.

[163] J. Tanenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reducton," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[164] J. Tarel, H. Civi, and D. Cooper, "Pose estimation of free-form 3D objects without point matching using algebraic surface models," in *Proc. of IEEE Workshop on Model-based 3D Image Analysis*, 1998, pp. 13–21.

[165] G. Taubin, "Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation," *IEEE Trans. Pattern Analysis and Machine Intell.*, vol. 13, no. 11, pp. 1115–1138, 1991.

[166] D. Tax and R. Duin, "Support vector data description," *Machine Learning*, vol. 54, pp. 45–66, 2004.

[167] J. Tian and K. Ma, "Markov Chain Monte Carlo super-resolution image reconstruction with simultaneous adaptation of the prior image model," in *Lecture Notes from Computer Science*, vol. 4261, Springer, 2006, p. 287.

[168] J. Tian and K. Ma, "A MCMC approach for Bayesian super-resolution image reconstruction," *Proc. IEEE International Conference on Image Processing*, vol. 1, pp. 45–48, 2005.

[169] R. Tibshirani, "Principal curves revisited," *Statistics and Computing*, vol. 2, pp. 183–190, 1992.

[170] P. Tondeur, *Foliations on Riemannian Manifolds*. New York: Springer-Verlag, 1988.

[171] D. Tuch, "Q-ball imaging," *Magnetic Resonance in Medicine*, vol. 52, no. 6, pp. 1358–1372, 2004.

[172] D. Tuch, T. Reese, M. Wiegell, N. Makris, J. Belliveau, and V. Wedeen, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magnetic Resonance in Medicine*, vol. 48, no. 4, pp. 577–582, 2002.

[173] U. Türe, M. Yasargil, A. Friedman, and O. Al-Mefty, "Fiber dissection technique: Lateral aspect of the brain," *Neurosurgery*, vol. 47, no. 2, p. 417, 2000.

[174] V. Vapnik, *Statistical Learning Theory. 1998*. Wiley, New York, 1998.

[175] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, New York, 1995.

[176] L. Velho and G. Taubin, "Estimating the in/out function of a surface represented by points," in *Proc. of the Eighth ACM Symposium on Solid Modeling and Applications*, ACM, 2003, pp. 108–114.

[177] C. Vidal, R. Nicolson, T. DeVito, K. Hayashi, J. Geaga, D. Drost, P. Williamson, N. Rajakumar, Y. Sui, and R. Dutton, "Mapping corpus callosum deficits in autism: an index of aberrant cortical connectivity," *Biological Psychiatry*, vol. 60, no. 3, pp. 218–225, 2006.

[178] G. Wahba, *Spline Models for Observational Data*. Society for Industrial Mathematics, 1990, vol. 59.

[179] M. Wahl, B. Lauterbach-Soon, E. Hattingen, P. Jung, O. Singer, S. Volz, J. Klein, H. Steinmetz, and U. Ziemann, "Human motor corpus callosum: Topography, somatotopy, and link between microstructure and function," *Journal of Neuroscience*, vol. 27, no. 45, p. 12 132, 2007.

[180] C. Walder, O. Chapelle, and B. Schölkopf, "Implicit Surface modelling as an eigenvalue problem," in *Proc. 22nd Int. Conf. on Machine Learning*, 2005, pp. 937–944.

[181] G. Wang, "Incorporating prior knowledge in Support Vector Machines: Retrospect and prospect," in *Proc. 4th Int. Conf. Networked Computing and Advanced Information Management*, vol. 1, 2008, pp. 464–468.

[182] R. Wang, T. Benner, A. Sorensen, and V. Wedeen, "Diffusion toolkit: A software package for diffusion imaging data processing and tractography," in *Proc. International Society for Magnetic Resonance in Medicine*, vol. 15, 2007, p. 3720.

[183] X. Wang, P. Tiňo, and M. Fardal, "Multiple manifold learning framework based on hierarchical mixture density model," in *Proc. European Conf. Machine Learning and Knowledge Discovery in Databases*, 2008, pp. 566–581.

[184] L. Wasserman, *All of Nonparametric Statistics*. Springer-Verlag, New York, 2006.

[185] ——, *All of Statistics: A Concise Course in Statistical Inference*. Springer Verlag, New York, 2004.

[186] V. Wedeen, D. Rosene, R. Wang, G. Dai, F. Mortazavi, P. Hagmann, J. Kaas, and W. Tseng, "The geometric structure of the brain fiber pathways," *Science*, vol. 335, no. 6076, pp. 1628–1634, 2012.

[187] K. Q. Weinberger, F. Sha, and L. K. Saul, "Learning a kernel matrix for nonlinear dimensionality reduction," in *Proc. 21st International Conference on Machine Learning*, 2004, pp. 839–846.

[188] E. Weiszfeld, *Sur le point pour lequel la somme des distances de n points donnés est minimum*. Tôhoku Imperial University, 1937.

[189] Y. Weldeselassie and G. Hamarneh, "DT-MRI segmentation using graph cuts," 31, vol. 8, Society of Photo-Optical Instrumentation Engineers, 2007, 65121K 1–9.

[190] Y. Weldeselassie, G. Hamarneh, M. Beg, and M. Atkins, "Novel decomposition of tensor distance into shape and orientation distances," in *Medical Image Computing and Computer-Assisted Intervention Workshop on Diffusion Modelling and the Fibre Cup*, 2009, pp. 173–180.

[191] C. Westin and H. Knutsson, "Tensor field regularization using normalized convolution," *Computer Aided Systems Theory-EUROCAST 2003*, pp. 564–572, 2003.

[192] C. Westin, M. Martin-Fernandez, C. Alberola-Lopez, J. Ruiz-Alzola, and H. Knutsson, "Tensor field regularization using normalized convolution and Markov random fields in a Bayesian framework," *Visualization and Image Processing of Tensor Fields.*, pp. 381–398, 2006.

[193] M. Wiegell, H. Larsson, and V. Wedeen, "Fiber crossing in human brain depicted with Diffusion Tensor MR Imaging," *Radiology*, vol. 217, no. 3, pp. 897–903, 2000.

[194] L. Williams and D. Jacobs, "Stochastic completion fields: A neural model of illusory contour shape and salience," *Neural Computation*, vol. 9, no. 4, pp. 837–858, 1997.

[195] R. Wolak, "Foliated and associated geometric structures on foliated manifolds," *Annales de la Faculté des Sciences de Toulouse. Série V. Mathématiques*, vol. 10, pp. 337–360, 1989.

[196] R. Woods, S. Grafton, C. Holmes, S. Cherry, and J. Mazziotta, "Automated image registration: I. General methods and intrasubject, intramodality validation," *Journal of Computer Assisted Tomography*, vol. 22, no. 1, pp. 139–152, 1998.

[197] J. Wu, X. Zhang, and L. Peng, "Positive approximation and interpolation using compactly supported radial basis functions," in *Mathematical Problems in Engineering*, Hindawi, vol. 2010, 2010, pp. 1–10.

[198] Q. Wu, "A correlation-relaxation-labeling framework for computing optical flow template matching from a new perspective," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 843–853, 1995.

[199] X. Wu and R. Srihari, "Incorporating prior knowledge with weighted margin Support Vector Machines," in *Proc. of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2004, pp. 326–333.

[200] C. Xu and J. Prince, "Generalized gradient vector flow external forces for active contours," *Signal Processing*, vol. 71, no. 2, pp. 131–139, 1998.

[201] S. Yeh, "Reconstruction of foliations from directional information," PhD thesis, University of St Andrews, 2007.

[202] L. Ying and E. Candès, "The phase flow method," *Journal of Computational Physics*, vol. 220, no. 1, pp. 184–215, 2006.

[203] P. Yushkevich, P. Thomas Fletcher, S. Joshi, A. Thall, and S. Pizer, "Continuous medial representations for geometric object modeling in 2D and 3D," *Image and Vision Computing*, vol. 21, no. 1, pp. 17–27, 2003.

[204] H. Zhang, B. Avants, P. Yushkevich, J. Woo, S. Wang, L. McCluskey, L. Elman, E. Melhem, and J. Gee, "High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis," *Transactions on Medical Imaging*, vol. 26, no. 11, pp. 1585–1597, 2007.

[205] J. Zhang, H. Ji, N. Kang, and N. Cao, "Fiber tractography in diffusion tensor magnetic resonance imaging: A survey and beyond," in *Proc. Int. Symp. on Medical Imaging and Computing*, 2005, pp. 39–56.

[206] K. Zhang and T. J. Sejnowski, "A universal scaling law between gray matter and white matter of cerebral cortex," *Proc. National Academy of Sciences*, vol. 97, no. 10, pp. 5621–5626, 2000.

[207] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, p. 45, 2001.

[208] H. Zhao, S. Osher, B. Merriman, and M. Kang, "Implicit and nonparametric shape reconstruction from unorganized data using a variational level set method," *Computer Vision and Image Understanding*, vol. 80, no. 3, pp. 295–314, 2000.

[209] X. Zhu and A. Goldberg, "Introduction to Semi-Supervised learning," *Synthesis lectures on artificial intelligence and machine learning*, vol. 3, no. 1, pp. 1–130, 2009.

[210] U. Ziyan, D. Tuch, and C. Westin, "Segmentation of thalamic nuclei from DTI using spectral clustering," *Proc. Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 807–814, 2006.