A Longitudinal Functional Analysis Framework for Analysis of White Matter Tract Statistics

Ying Yuan[1], John H. Gilmore[2], Xiujuan Geng[2], Martin A. Styner[2,3], Kehui Chen[4], Jane-Ling Wang[5], and Hongtu Zhu[6,7]

[1] Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN 38105, USA
[2] Department of Psychiatry, and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[3] Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[4] Department of Statistics, University of Pittsburgh, PA 15260, USA
[5] Department of Statistics, University of California at Davis,CA 95616, USA
[6] Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[7] Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Abstract. Many longitudinal imaging studies have been/are being widely conducted to use diffusion tensor imaging (DTI) to better understand white matter maturation in normal controls and diseased subjects. There is an urgent demand for the development of statistical methods for analyzing diffusion properties along major fiber tracts obtained from longitudinal DTI studies. Jointly analyzing fiber-tract diffusion properties and covariates from longitudinal studies raises several major challenges including (i) infinite-dimensional functional response data, (ii) complex spatial-temporal correlation structure, and (iii) complex spatial smoothness. To address these challenges, this article is to develop a longitudinal functional analysis framework (LFAF) to delineate the dynamic changes of diffusion properties along major fiber tracts and their association with a set of covariates of interest (e.g., age and group status) and the structure of the variability of these white matter tract properties in various longitudinal studies. Our LFAF consists of a functional mixed effects model for addressing all three challenges, an efficient method for spatially smoothing varying coefficient functions, an estimation method for estimating the spatial-temporal correlation structure, a test procedure with a global test statistic for testing hypotheses of interest associated with functional response, and a simultaneous confidence band for quantifying the uncertainty in the estimated coefficient functions. Simulated data are used to evaluate the finite sample performance of LFAF and to demonstrate that LFAF significantly outperforms a voxel-wise mixed model method. We apply LFAF to study the spatial-temporal dynamics of white-matter fiber tracts in a clinical study of neurodevelopment.

J.C. Gee et al. (Eds.): IPMI 2013, LNCS 7917, pp. 220–231, 2013.

1 Introduction

Many DTI studies have specifically focused on the region-of-interest (ROI) and voxel analyses of diffusion properties, such as fractional anisotropy (FA), to delineate white matter maturation and integrity along major fiber tracts in cross-sectional studies [1], [2], [3], [4], [5]. As discussed in [6], [7], [5], standard ROI and voxel-based analyses suffer from serious drawbacks, such as poor alignment quality and the identification of meaningful ROIs. To address these drawbacks, developing fiber-tract based analysis of diffusion properties have received much more attention recently [7], [8], [9], [6], [4], since major fiber tracts are much more objective, specific, and reliable ROIs compared with anatomically defined ROIs. Statistically, several functional regression models have been developed to analyze fiber-tract diffusion properties and covariates from cross-sectional studies [4], [5]. However, cross-sectional studies have limited power in characterizing individual white matter maturation compared with longitudinal studies.

Longitudinal DTI studies have been used to characterize individual change in diffusion properties over time and the effect of some covariates, such as gender, on the individual change in different age groups [10], [1], [11], [12]. A distinctive feature of longitudinal data is that measurements of the same individual usually exhibit positive correlation and the strength of the correlation decreases with the time separation. Ignoring temporal correlation structure in measures would likely influence subsequent statistical inference, such as increase in false positive and negative errors, which may lead to misleading scientific inference [13], [14]. Recently, linear and nonlinear mixed effects models have been used to explicitly account for the temporal correlation in the ROI analysis of longitudinal diffusion properties [14], [11], [12]. Moreover, in [1], a functional mixed effects model proposed by [15] is used to analyze fiber-tract diffusion properties from longitudinal studies. However, since the original mixed effects model in [15] was developed to model the data with functional responses(of time or distance) measured once for each subject, it is inappropriate to directly apply the method in [15] to longitudinally measured fiber-tract diffusion properties, which are essentially infinite-dimensional functional data measured across multiple time points instead of only one time point. These mixed effect models cannot be used to jointly analyze fiber-tract based diffusion properties and covariates from longitudinal studies due to at least three major challenges including (i) infinite-dimensional functional response data measured across time, (ii)complex spatial-temporal correlation structure, and (iii) complex spatial smoothness. According to the best of our knowledge, little has been done on the development of advanced statistical methods to address these challenges.

To fill up such gaps, we develop LFAF for the joint analysis of fiber-tract diffusion properties and covariates from longitudinal studies. Specifically, we propose a functional mixed effect model with two components: a multivariate varying-coefficient model for characterizing the dynamic association between fiber-tract diffusion properties and some covariates and a set of functional random effects for capturing complex spatial-temporal correlation structure. Our LFAF is closely related with a longitudinal functional principal component analysis in [16], but

we make three major advances in formal statistical inference as follows. The first one is to develop an efficient estimation method to spatially smooth varying coefficient functions, while accounting for spatial-temporal correlation structure. The second one is to propose a test procedure with a global test statistic for testing hypotheses of interest associated with functional response. The third one is to approximate a simultaneous confidence band for quantifying the uncertainty in the estimated coefficient functions. LFAF provides a rigorous analytical tool for characterizing the dynamic changes of functional response data and their association with a set of covariates.

2 Methodologies

To compare fiber-tract diffusion properties across subjects and populations, we use DTI atlas building with a group-wise longitudinal large deformation diffeomorphic registration method followed by atlas fiber tractography and fiber parametrization as described in [1], [6] to extract DTI fibers and establish DTI fiber correspondence across all DTI datasets from different subjects at all time points. Since this method has been described in detail in [1], we do not include them here for the sake of space.

The aim of this paper is to present a longitudinal functional analysis pipeline for delineating the dynamic changes of fiber tract diffusion properties and their association with a set of covariates of interest, such as age. We discuss each step of LFAF below and will present the asymptotic properties of all estimators and the test statistic with detailed assumptions and proofs elsewhere.

Functional Mixed Effects Model. Consider a longitudinal study with n independent subjects. Let $s \in [0, L]$ be the arc length of any point on a specific fiber tract relative to a fixed end point of the fiber tract and $\{s_m : m = 1, \ldots, M\}$ a set of M grid points in $[0, L]$, where L is the longest arc length on the fiber tract. Let $y_{ij}(s)$ be a specific diffusion property measured at s and $\mathbf{x}_{ij}$ be a $p_x \times 1$ vector of covariates measured at time t_{ij} for the i-th subject for $j = 1, \cdots, r_i$ where r_i is the total number of time points for the subject i. A functional mixed model is given by

$$y_{ij}(s) = \mathbf{x}_{ij}^T B(s) + \mathbf{z}_{ij}^T \boldsymbol{\xi}_i(s) + \eta_{ij}(s) + \epsilon_{ij}(s), \tag{1}$$

where $\boldsymbol{\xi}_i(s)$ is $p_z \times 1$ vector of functional random effects, $\mathbf{z}_{ij}$ is a $p_z \times 1$ vector of covariates associated with $\xi_i(s)$ and commonly a subvector of $\mathbf{x}_{ij}$, $\eta_{ij}(s)$ is a random function for subject i at time t_{ij}, $\epsilon_{ij}(s)$ is a measurement error, and $B(s) = (\beta_1(s), \ldots, \beta_{p_x}(s))^T$ is a $p_x \times 1$ vector of functions of s. Moreover, $\boldsymbol{\xi}_i(s)$ primarily characterizes the temporal variations and within-subject dependence, while $\eta_{ij}(s)$ characterizes individual curve variations from $\mathbf{x}_{ij}^T B(s) + \mathbf{z}_{ij}^T \boldsymbol{\xi}_i(s)$ and spatial dependence. It is also assumed that $\boldsymbol{\xi}_i(s)$, $\boldsymbol{\epsilon}_i(s) = (\epsilon_{i1}(s), \cdots, \epsilon_{ir_i}(s))^T$, and $\boldsymbol{\eta}_i(s) = (\eta_{i1}(s), \cdots, \eta_{ir_i}(s))^T$ are mutually independent and identical copies of $\mathrm{SP}(\mathbf{0}, \Sigma_\xi)$, $\mathrm{SP}(\mathbf{0}, \Sigma_\eta)$ and $\mathrm{SP}(\mu, \Sigma_\epsilon)$, respectively, where $\mathrm{SP}(\mu, \Sigma)$ denotes a stochastic process vector with mean function $\mu(s)$ and covariance function

$\Sigma(s,t)$. Moreover, $\boldsymbol{\epsilon}_i(s)$ and $\boldsymbol{\epsilon}_i(t)$ are assumed to be independent for $s \neq t$ and $\Sigma_\epsilon(s,t)$ takes the form of $\sigma_\epsilon(s)^2\mathbf{1}(s=t)$, where $\mathbf{1}(\cdot)$ is an indicator function.

Initial Estimator of Varying Coefficient Functions. We use the local linear regression method and the weighted least squares estimation to estimate $B(s)$ [17]. Since the local linear regression method adapts automatically at the boundary points [17], it is ideal for dealing with scalar diffusion properties along fiber tracts with two ends. Specifically, for a specific bandwidth h, we estimate $B(s)$ by minimizing the following weighted least squares function

$$\sum_{i=1}^{n}\sum_{j=1}^{r_i}\sum_{m=1}^{M}\{y_{ij}(s_m) - \mathbf{x}_{ij}^T[B(s) + \dot{B}(s)(s_m - s)]\}^2 K_h(s_m - s),$$

where $K(s)$ is a kernel function and $K_h(s) = h^{-1}K(s/h)$. The optimal estimator of $B_j(s)$, denoted by $\hat{B}_j(s)$ is obtained at the optimal bandwidth selected by using a leave-one-curve-out cross-validation method.

Estimating Covariance Functions. The covariance structure of $y_i(s)$ plays a crucial role in our proposed inference procedure. We proposed the following estimation procedure to estimate the covariance functions in model (1).

- (I)Use the local constant regression method to estimate $\Sigma_\xi(s,t)$ and $\Sigma_\xi(s,t)$ for each pair $s \leq t$ [17], which yields the estimates $\tilde{\Sigma}_\xi(s,t)$ and $\tilde{\Sigma}_\eta(s,t)$.
- (II) Use the local constant regression method to smooth $\tilde{\Sigma}_\xi(s,t)$ to calculate $\hat{\Sigma}_\xi(s,t)$ across (s,t) and to smooth $\tilde{\Sigma}_\eta(s,t)$ for $s \neq t$ to calculate $\hat{\Sigma}_\eta(s,t)$ across (s,t).

The above procedure cannot guarantee that $\hat{\Sigma}_\xi(s,t)$ and $\hat{\Sigma}_\eta(s,t)$ are semipositive definite. We apply an adjustment procedure in [18] to transform $\hat{\Sigma}_\xi(s,t)$ and $\hat{\Sigma}_\eta(s,t)$ into semipositive definite covariance functions. Its key idea is to approximate the covariance functions by truncating the components corresponding to negative eigenvalues in the spectral representations of $\hat{\Sigma}_\xi(s,t)$ and $\hat{\Sigma}_\eta(s,t)$.

Refined Estimator of Varying Coefficient Functions. In the initial estimation of varying coefficient functions, it is assumed that there are no spatial-temporal correlations. At this stage, we refine the estimate of $B(\cdot)$ by incorporating the estimated spatial-temporal covariance functions for each subject. Let $\mathbf{X}_i = [\mathbf{x}_{i1}\cdots\mathbf{x}_{ir_i}]$ be a $p_x \times r_i$ matrix. We reestimate $B(s)$ by minimizing the following weighted least squares function

$$\sum_{i=1}^{n}\sum_{m=1}^{M}\{[\mathbf{y}_i(s_m) - \mathbf{X}_i^T(B(s) + \dot{B}(s)(s_m - s))]^T\hat{\Sigma}_{y,i}^{(1)}(s_m,s_m)^{-1/2}\}^{\otimes 2}K_h(s_m - s),$$

where $\hat{\Sigma}_{y,i}^{(1)}(s,t) = \mathbf{z}_{ij}^T\hat{\Sigma}_\xi(s,t)\mathbf{z}_{ij} + \hat{\Sigma}_\eta(s,t)$. We also select the bandwidth h by using the leave-one-curve-out cross-validation method.

Smoothing Individual Functions. Let $g_{ij}(s) = \mathbf{z}_{ij}^T \boldsymbol{\xi}_i(s) + \eta_{ij}(s)$ for all i, j. We also employ the local linear regression technique to estimate all individual functions $g_{ij}(s)$ [17]. Specifically, we estimate $g_{ij}(s)$ by minimizing the weighted least squares function given by

$$\sum_{m=1}^{M} [y_{ij}(s_m) - \mathbf{x}_{ij}^T \hat{B}(s_m) - (g_{ij}(s) + \dot{g}_{ij}(s)(s_m - s))]^2 K_h(s_m - s).$$

The optimal estimator, $\hat{g}_{ij}(s)$ is obtained at the optimal bandwidth selected by using the leave-one-out generalized cross-validation method.

Hypothesis Test. We test the null hypothesis $H_0 : CB(s) = \mathbf{b}_0(s)$ for all s against $H_1 : CB(s) \neq \mathbf{b}_0(s)$ by proposing a global test statistics $\mathbf{S}_n$, defined by

$$\mathbf{S}_n = \int_0^1 \mathbf{d}(s)^T \{C[\sum_{i=1}^{n} \mathbf{X}_i \hat{\Sigma}_{y,i}^{(1)}(s,s)^{-1} \mathbf{X}_i^T]^{-1} C^T\}^{-1} \mathbf{d}(s) ds, \tag{2}$$

where $\mathbf{d}(s) = C\text{vec}(\hat{B}(s) - \text{bias}(\hat{B}(s))) - \mathbf{b}_0(s)$. Following [19], a smaller bandwidth leads to a smaller value of bias$(\hat{B}(s))$. Moreover, according to our simulation studies below, we have found that the effect of dropping bias$(\hat{B}(s))$ is negligible, and therefore we drop it from now on. The asymptotic distribution of $\mathbf{S}_n$ is very complicated and it is difficult to directly approximate the percentiles of $\mathbf{S}_n$ under the null hypothesis. Instead, we propose to use a resampling method to approximate the p-value of $\mathbf{S}_n$ as in [20], [5].

Simultaneous Confidence Bands. For a given significance level α, we construct a simultaneous confidence band for each $\beta_l(s)$ such that

$$P(\hat{\beta}_l^{L,\alpha}(s) < \beta_l(s) < \hat{\beta}_l^{U,\alpha}(s) \text{ for all } s \in [0,1]) = 1 - \alpha, \tag{3}$$

where $\beta_l^{L,\alpha}(s)$ and $\beta_l^{U,\alpha}(s)$ are the lower and upper limits of the confidence band. We develop a resampling method to approximate $C_l(\alpha)$ as in [20], [5].

3 A Real Example and Simulation Studies

A Real Example. We applied LFAF to investigate neonatal brain development in a longitudinal study consisting of 137 healthy infants (83 males and 54 females) with mean age at the baseline 297.89 days with SD: 13.9 days scanned at 2 week, 1 year, and 2 year. A 3T Allegra head-only MR system (Siemens Medical Solutions, Erlangen, Germany) was used to acquire all the images. The system was equipped with a maximal gradient strength of 40 mT/m and a maximal slew rate of 400 mT/(m · ms). The DTI images were obtained by using a single shot EPI DTI sequence with the following variables: TR/TE= 5200/73 ms, slice thickness= 2 mm, in-plane resolution= 2×2 mm^2 with eddy current compensation. In this study, two sequences with 6 or 42 non-collinear directions,

respectively, were applied at the b-value of 1000 s/mm^2 with a reference scan ($b = 0$). When the sequence with 6 gradient directions was applied, to improve the signal-to-noise ratio of the images, a total of five scans were acquired and averaged. A weighted least squares estimation method were used to estimate diffusion tensors [21]. Then a DTI atlas building with a group-wise longitudinal large deformation diffeomorphic registration method followed by atlas fiber tractography and fiber parametrization as described in [1], [6] were used to extract DTI fibers and establish DTI fiber correspondence across all DTI datasets from different subjects at all time points. We chose the genu of the corpus callosum to illustrate our LFAF. Three diffusion properties were extracted along the selected fiber tracts including FA, RD, and AD, at each grid point and each time point for all 137 infants.

In this study, we have three specific aims. The first is to investigate the gender effect on the development of the fiber tract diffusion properties. The second is to investigate the number of gradient directions effect on the the fiber tract diffusion properties. The third is to delineate the development of fiber tract diffusion properties over time. As a graphical illustration, fractional anisotropy (FA), axial diffusivity (AD) and radial diffusivity (RD) values were plotted along the genu of the corpus callosum for all subjects within each age group (Fig. 1 (a)-(c)). They were also plotted for 35 selected subjects at a selected grid point (Fig. 1 (d)-(f)). An obvious increasing trend for the values of FA, and obvious decreasing trends for the values of RD and AD were observed at nearly all grid points, especially from neonate to the first year. It is also observed from Fig. 1 (a)-(c) that there is a random subject-to-subject variation in FA, RD and AD measures at each grid point along the two tracts. In addition, Fig. 1 (d)-(f) shows that there is a random subject-to-subject variation in the age effect on FA, RD and AD measures at the selected location. Thus, we use our functional mixed model to fit this dataset and to statistically test the gender, number of gradient directions and age effects on FA, RD and AD values along this tract.

We fit model (1) to the fiber-tract FA, RD and AD values from all 137 subjects, in which $\mathbf{x} = (1, \mathrm{Dir}, \mathrm{G}, \mathrm{Age}_1, \mathrm{Age}_2)^T$, $\mathbf{z} = (1, \mathrm{Age}_1, \mathrm{Age}_2)^T$ and Age_1 (or Age_2) is an indicator variable indicating whether a subject belongs to the first (or second) year age group. The coefficient functions related to Age_1 and Age_2 can be used to investigate whether there is some change from neonate to the first year of life, from the first year to the second year and from neonate to the second year. For the hypothesis testing, we constructed the global test statistic $\mathbf{S}_n$ to test the gender, number of gradient directions and age effects on FA, RD and AD values. We approximated the p-value of $\mathbf{S}_n$ using the resampling method with $G = 5,000$ replications. We also constructed the 95% simultaneous confidence bands for the functional coefficients $B(s)$.

Fig. 2 presents the estimated coefficient functions $B(s)$. The intercept functions describe the overall trend of the three diffusion properties. The coefficient functions of Age_1 and Age_2 for FA (the fourth and fifth rows in Fig. 2 (a)) are positive at all grid points, which indicates that FA increases from neonate to the first year and from neonate to the second year along the genu of the corpus

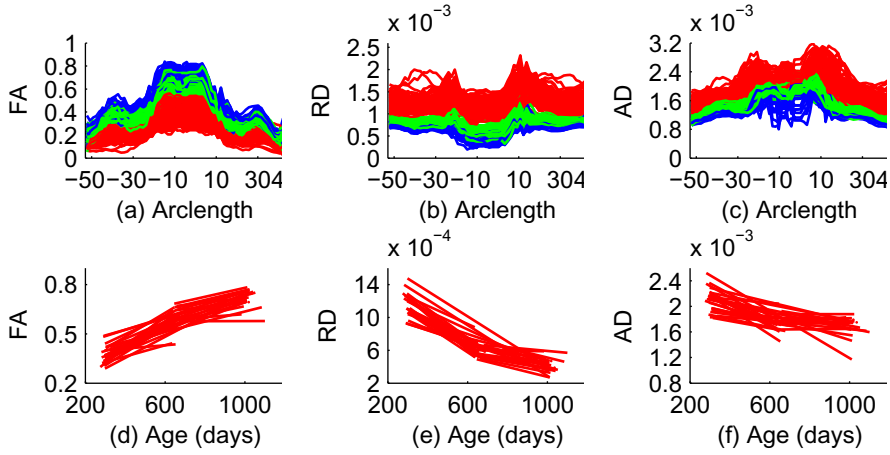

Fig. 1. FA (a), RD (b) and AD (c) values along the genu of the corpus callosum for all 137 subjects in each age group,with red, green and blue colors for neonate, one-year-old and two-year-old infants, separately. FA (d), RD (e) and AD (f) values varying across age at a selected location (arclength=-8.75) for the 35 selected subjects.

callosum, while the coefficients of Age_1 and Age_2 for RD and AD (the fourth and fifth rows in Fig. 2 (b) and (c)) are negative at all grid points, which indicates that RD and AD decrease from neonate to the first year and from neonate to the second year .

It is observed that the coefficients related to the number of gradient directions for FA (the second panel in Fig. 2 (a)) are positive in the central region of the tract, which indicates that compared with FA values obtained by DTI with 6 gradient directions, FA values with 42 gradient directions are larger in the central region of the tract while The coefficients of Dir for RD (the second panel in Fig. 2 (b)) are negative in the central regions, which indicates that compared with RD values obtained by DTI with 42 gradient directions, RD values with 6 gradient directions are larger in the central region of the tract.

Fig. 3 presents mean diffusion profiles along the tract at each age group. FA and AD values in the central region of the genu of the corpus callosum are larger compared to peripheral regions. RD values in the region close to mid-sagittal brain is smaller. They are consistent with general topological rules of white matter maturation as pointed in [1].

The hypothesis testing results show that there are significant age and number of gradient direction effects on FA, RD and AD values, which agrees with the findings in Fig. 2 (a)-(c). FA's, RD's and AD's are significantly different between neonate versus the first year, and between the first year versus the second year with p value $< .0001$, far smaller than 0.05 significance level. It is observed from Fig. 3 that mean FA values increase from neonate to the first year and then from the first year to the second year while mean RD and AD decreases from neonate to the first year and then from the first year to the second year at any

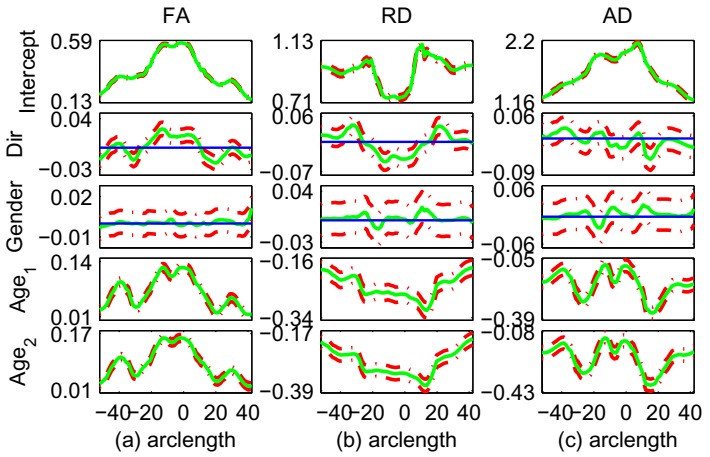

Fig. 2. 95% simultaneous confidence bands for varying coefficient functions for FA (a), RD (b) and AD (c) along the genu of the corpus callosum. The solid curves are the estimated coefficient functions, and the dashed curves are the 95% confidence bands. The thin horizontal line is the line crossing the origin.

grid point along the tract. Moreover, the change from the neonate to the first year is larger than that from the first year to the second year. In addition, the number of gradient directions was found to have the significant effect on FA, RD and AD values. It is shown in Fig. 3 that in general, FA values are smaller while RD and AD values are larger when 6 gradient directions instead of 42 gradient directions is used, especially in the central regions. No gender differences in FA, RD and AD were found for this tract.

Simulation Studies. We conducted a Monte Carlo simulation study to examine the finite sample performance of LFAF . At each point s_m along the genu, the noisy FA's are simulated according to the following model:

$$y_{ij}(s_m) = \mathbf{x}_{ij}^T B(s_m) + \tau_i g_{ij}(s_m) + \tau_{ij}(s_m)\epsilon_{ij}(s_m) \tag{4}$$

where τ_i and $\tau_{ij}(s_m)$ were independently generated from $N(0,1)$ random generators for $i = 1, \cdots, n$, $j = 1, \cdots, r_i$, and $m = 1, \cdots, M$. Specifically, we set $n = 137$, $M = 64$, $x_{ij} = (1, Dir_{ij}, G_i, Age_{ij})^T$ and $z_{ij} = (1, Age_{ij})^T$, where Dir_{ij}, G_i and Age_{ij}, respectively, denote the indicator of different numbers of gradient directions used, gender and age at MRI scanning. To mimic real imaging data, we applied our proposed LFAF method to FA's along the genu from all 137 infants in our clinical data to estimate $B(s)$ by $\hat{B}(s)$, $g_{ij}(s)$ by $\hat{g}_{ij}(s)$, and $\epsilon_{ij}(s)$ by $\hat{\epsilon}_{ij}(s)$. According to our real data analysis, the age effect is significant for our clinical data. So we fixed all functions in $B(s)$ at their corresponding functions in $\hat{B}(s)$ except that the fourth function of $B(s)$, denoted by $\beta_4(s)$, was set as c times the fourth column of $\beta_4(s)$ where c is set at different values in

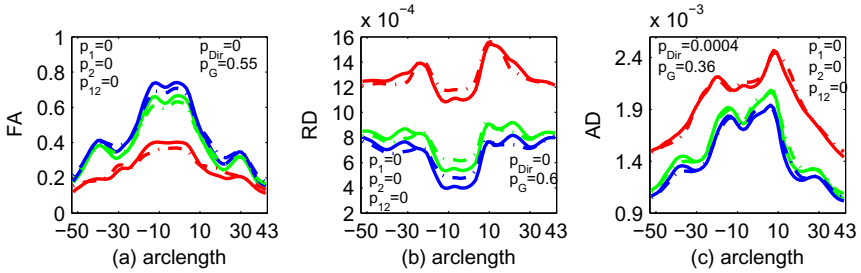

Fig. 3. Mean FA (a), RD (b) and AD (c) values along the genu of the corpus callosum and the gender, number of gradient directions and age effects on FA, RD and AD values. Solid lines are mean FA, RD and AD curves for the 42 gradient direction while the dashed lines are for the six gradient directions. Red, green and blue colors for neonate, one-year-old and two-year-old infants, separately. p_1 is the p value for the difference in the diffusion measure between neonate and the first year, p_2 is the p value for the difference in the diffusion measure between neonate and the second year, p_{12} is the p value for the difference between the first year and the second year, p_G is the p value for the gender effect, p_{Dir} is the p value for the effect of the number of gradient directions.

order to study the Type I and II error rates of our global test statistic in testing the age effect.

We have three aims in this simulation study. The first aim is to examine the coverage probabilities of the simultaneous confidence bands for all varying coefficient functions $\beta_l(s)$ for $l = 1, 2, 3, 4$ in $B(s)$. We only considered the simulated FA measures at $c = 0.1$ and constructed the 95% and 99% simultaneous confidence bands for all $\beta_l(s)$. Table 1 summarizes the empirical coverage probabilities based on 500 replications for $\alpha = 0.01$ and 0.05. The coverage probabilities are quite close to the prespecified confidence levels.

Table 1. Simulated coverage probabilities for varying coefficient functions in $B(s) = (\beta_l(s))$ based on 500 replications at the significance levels $\alpha = 0.01$ and 0.05

	$\alpha = 0.01$				$\alpha = 0.05$			
c	intercept	Dir	Gender	Age	intercept	Dir	Gender	Age
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 1$	$l = 2$	$l = 3$	$l = 4$
0.1	0.942	0.930	0.946	0.946	0.992	0.986	0.986	0.980

The second aim is to evaluate the Types I and II error rates of the global test statistic $\mathbf{S}_n$. In neuroimaging studies, some scientific questions require the assessment of the development of white matter across age. We formulated the questions as testing the null hypothesis $H_0 : \beta_4(s) = 0$ for all s along the genu against $H_1 : \beta_4(s) \neq 0$ for at least one s on the tract. We first fixed $c = 0$ to assess the Type I error rates for $\mathbf{S}_n$, and then set $c = 0.02, 0.04, 0.06, 0.08$, and

0.1 to examine the Type II error rates for $\mathbf{S}_n$ at different effect sizes. In order to evaluate the Types I and II error rates at different sample sizes, we let $n = 137$ and 70. For $n = 137$, the values of indicator of number of gradient directions, gender and age were set the same as the 137 subjects in our clinical study. For $n = 70$, we randomly chose 35 males and 35 females from the 137 subjects and used their values for indicator of number of gradient directions, gender and age to simulate the values of FA along the genu tract. We applied LFAF to the simulated FA measures along the genu and approximated the p-value of $\mathbf{S}_n$ by using the resampling method with $G = 500$. For each c, we set the significance level α at both 0.05 and 0.01 and used 500 replications to estimate the rejection rate of $\mathbf{S}_n$. At a fixed α, if the Type I rejection rate is smaller than α, then the test is conservative, whereas if the Type I rejection rate is greater than α, then the test is anticonservative, or liberal. Fig. 4 presents the rejection rates of $\mathbf{S}_n$ across all effect sizes at the two significance levels ($\alpha = 0.05$ or 0.01). It is observed that Type I error rates are well maintained with the values 0.048 and 0.01, respectively, at the two significance levels. In addition, the statistical power for rejecting the null hypothesis increases with the sample size, the effect size and the significance level, which is consistent with our expectation.

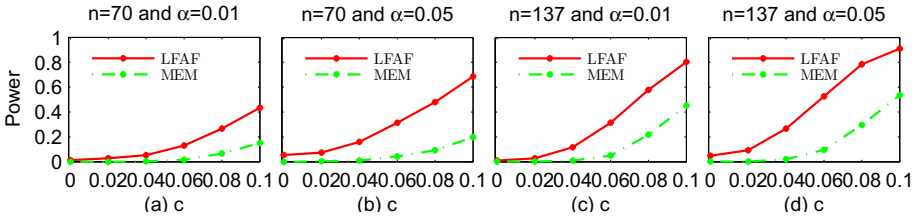

Fig. 4. Simulation study: Type I and Type II error rates as functions of c. Rejection rates of $\mathbf{S}_n$ based on the resampling method are calculated at six different values of the effect size c for sample size 70 and 137 at the 0.01 and 0.05 significance levels using LFAF and MEM.

The third aim is to show that LFAF outperforms mixed model. To this end, we first fitted linear mixed model at each s_m without separating $\eta_{ij}(s_m)$ from $\epsilon_{ij}(s_m)$. Then, we calculated the global testing statistic and approximated the p-values with the resampling method as in LFAF. Fig. 4 shows that the linear mixed model is much less powerful than LFAF.

4 Discussion

We presented a rigorous and efficient statistical method for analyzing longitudinally measured diffusion tensor properties along major white matter fiber tracts and showed that this method has the greater statistical power in detecting the effects of covariates of interest as is shown in Fig. 4 and the greater accuracy

in characterizing the uncertainty in estimating coefficient functions as in Table 1. The proposed method was successfully applied to study the normal growth of white matter fiber tracts in the clinical study of neurodevelopment by revealing the complex spatiotemporal maturation patterns as the apparent changes in fiber tract diffusion properties. It can also be used to evaluate the time course of white matter disease induced by the treatment, such as high-dose chemotherapy, in clinical trials. Besides, because of its greater statistical power than voxelwise mixed effects methods, more significant findings could be obtained in the genome wide studies of the associations between the genetic markers with the white matter fiber tract phenotypes by using LFAF.

Acknowledgements. This work was partially supported by NIH grants RO1ES17240, MH091645, U54 EB005149, P30 HD03110, RR025747-01, P01CA142538-01, MH086633, MH064065, HD053000, and MH070890 to Dr. Zhu, NSF DMS 12-28369 and NSF DMS 09-06813 to Prof. Wang.

References

1. Geng, X., Gouttard, S., Sharma, A., Gu, H., Styner, M., Lin, W.: Quantitative tract-based white matter development from birth to age 2 years. NeuroImage 61, 542–557 (2012)
2. Ding, X.Q., Sun, Y., Braass, H., Illies, T., Zeumer, H., Lanfermann, H., Fiehler, J.: Evidence of rapid ongoing brain development beyond 2 years of age detected by fiber tracking. American Journal of Neuroradiology 29, 1261–1265 (2008)
3. Agosta, F., Henry, R.G., Migliaccio, R., Neuhaus, J., Miller, B.L., Dronkers, N.F., Brambati, S.M., Filippi, M., Ogar, J.M., Wilson, S.M., Gorno-Tempini, M.L.: Language networks in semantic dementia. Brain 133, 286–299 (2010)
4. Zhu, H., Styner, M., Tang, N., Liu, Z., Lin, W., Gilmore, J.H.: Frats: Functional regression analysis of dti tract statistics. IEEE Trans. Med. Imaging 29, 1039–1049 (2010)
5. Zhu, H., Kong, L., Li, R., Styner, M., Gerig, G., Lin, W., Gilmore, J.H.: Fadtts: Functional analysis of diffusion tensor tract statistics. NeuroImage 56, 1412–1425 (2011)
6. Goodlett, C.B., Fletcher, P.T., Gilmore, J.H., Gerig, G.: Group analysis of dti fiber tract statistics with application to neurodevelopment. NeuroImage 142, S133–S142 (2009)
7. Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.: Tractbased spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage 31, 1487–1505 (2006)
8. O'Donnell, L.J., Westin, C.-F., Golby, A.J.: Tract-based morphometry for white matter group analysis. Neuroimage 45, 832–844 (2009)
9. Yushkevich, P.A., Zhang, H., Simon, T.J., Gee, J.C.: Structure-specific statistical mapping of white matter tracts. Neuroimage 41, 448–461 (2008)
10. Evans, A.C., Brain Development Cooperative Group: The nih mri study of normal brain development. NeuroImage 30, 184–202 (2006)

11. Lebel, C., Beaulieu, C.: Longitudinal development of human brain wiring continues from childhood into adulthood. The Journal of Neuroscience 31, 10937–10947 (2011)
12. Sadeghi, N., Prastawa, M., Fletcher, P.T., Wolff, J., Gilmore, J.H., Gerig, G.: Regional characterization of longitudinal dt-mri to study white matter maturation of the early developing brain. NeuroImage (in press, 2013)
13. Diggle, P., Heagerty, P., Liang, K.Y., Zeger, S.: Analysis of Longitudinal Data, 2nd edn. Oxford University Press, New York (2002)
14. Fitzmaurice, G.M., Laird, N.M., Ware, J.H.: Applied Longitudinal Analysis. Wiley, New York (2004)
15. Guo, W.: Functional mixed effects models. Biometrics 58, 121–128 (2002)
16. Greven, S., Crainiceanu, C., Caffo, B., Reich, D.: Longitudinal functional principal component analysis. Electron. J. Statist. 4, 1022–1054 (2010)
17. Wand, M.P., Jones, M.C.: Kernel Smoothing. Chapman and Hall, London (1995)
18. Hall, P., Muller, H.G., Yao, F.: Modelling sparse generalized longitudinal observations with latent gaussian processes. J. R. Statist. Soc. B 70, 703–723 (2008)
19. Fan, J., Zhang, W.: Simultaneous confidence bands and hypothesis testing in varying-coefficient models. Scand. J. Statist. 27(4), 715–731 (2000)
20. van der Vaar, A.W., Wellner, J.A.: Weak Convergence and Empirical Processes. Springer-Verlag Inc. (1996)
21. Basser, P.J., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. Biophysical Journal 66, 259–267 (1994)