



A chi-square type test for time-invariant fiber pathways of the brain

Juna Goo¹ · Lyudmila Sakhanenko² · David C. Zhu³

Received: 29 December 2020 / Accepted: 11 January 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

A longitudinal diffusion tensor imaging (DTI) study on a single brain can be remarkably useful to probe white matter fiber connectivity that may or may not be stable over time. We consider a novel testing problem where the null hypothesis states that the trajectories of a coherently oriented fiber population remain the same over a fixed period of time. Compared to other applications that use changes in DTI scalar metrics over time, our test is focused on the partial derivative of the continuous ensemble of fiber trajectories with respect to time. The test statistic is shown to have the limiting chi-square distribution under the null hypothesis. The power of the test is demonstrated using Monte Carlo simulations based on both the theoretical and empirical critical values. The proposed method is applied to a longitudinal DTI study of a normal brain.

Keywords Functional central limit theorem · Nadaraya–Watson type kernel estimator · White matter fiber tractography

1 Introduction

Diffusion weighted imaging (DWI) is an established magnetic resonance imaging (MRI) technique for early diagnosis and prognosis of brain disorders that have axonal damage, such as Alzheimer’s disease (Zhu et al. 2013). It is not only supported by the advantages of conventional MRI such as its radiation-free imaging ability, but also reflects the diffusion of water molecules in the tissue of the brain when generating images. Diffusion tensor imaging (DTI) is a modeling method to reconstruct white matter fiber tracts based on the DWI data which includes multiple magnetic field gradient direction changes to be sensitive to directional water diffusion. In the white matter of the brain, a coherently oriented fiber bundle hinders the diffusion of water molecules in the directions perpendicular to the fiber population. Thus, water molecules diffuse anisotropically and the movement results in a 3×3 symmetric and

✉ Juna Goo
junagoo@boisestate.edu

¹ Department of Mathematics, Boise State University, Boise, ID, USA

² Department of Statistics and Probability, Michigan State University, East Lansing, MI, USA

³ Department of Radiology, Michigan State University, East Lansing, MI, USA

positive definite second-order tensor that is known as a diffusion tensor in DTI. The diffusion tensor is fully determined by six distinct elements due to its symmetry property.

In DTI-based fiber tractography, the magnitudes and directions of the diffusion tensor are of great importance to map neuronal structural connectivity. Each positive eigenvalue quantifies the diffusivity of the diffusion tensor in the corresponding direction of the mutually orthogonal eigenvectors. In particular, the dominant eigenvector of the diffusion tensor associated with the largest eigenvalue reveals the orientation of the coherent fiber population in a region of interest (ROI). The fundamental physics of DTI and its application to fiber tractography can be found in Johansen-Berg and Behrens (2014), Jones (2011), Mori and Tournier (2013), and references therein. Over the last two decades, many longitudinal DTI studies have employed existing statistical methods to examine longitudinal changes in DTI measurements using metrics such as fractional anisotropy (the normalized standard deviation of eigenvalues), mean diffusivity (the average of eigenvalues), axial diffusivity (the largest eigenvalue), and radial diffusivity (the average of the two smaller eigenvalues) over time (Dinkel et al. 2014; Kumar et al. 2012; Shaffer et al. 2017; Wu et al. 2010). In particular, parametric methods such as generalized linear models are used on those metrics.

However, studying the time-dependent change in the continuous ensemble of fiber trajectories requires a different approach that harnesses the whole brain information in the region that surrounds the fiber. While current methods based on eigenvalues of the diffusion tensor at different time points convey very partial information of the anatomy of the brain, our approach fully takes advantage of the whole brain in a longitudinal manner. We primarily focus on a hypothesis test for the problem of detecting the rate of change in the trajectories of the coherently oriented fiber population over time in a certain ROI. The proposed test is motivated by the need to detect the subtle but critical longitudinal changes in the orientation of fiber trajectories that may be caused by normal aging and neurodegenerative diseases. Previous similar works by Koltchinskii et al. (2007) and Carmichael and Sakhanenko (2016) modeled and studied DTI-based fiber tractography in non-longitudinal setting.

The rest of the paper consists of the following sections. We describe the mathematical framework of the testing problem in Sect. 2, followed by the estimation procedure in Sect. 3. Two main theorems are given in Sect. 4. Numerical approximation is briefly illustrated in Sect. 5. Monte Carlo simulations and real data analysis are presented in Sects. 6 and 7, respectively. Future research directions arising from some limitations are addressed in Sect. 8.

2 Framework

Let $x = [x_1 \ x_2 \ x_3]^\top \in \mathcal{X}$, where $\mathcal{X}$ is a compact set in $\mathbb{R}^3$. Let $t \in [0, T]$ with $T > 0$ in $\mathbb{R}$. Let $u = (x, t) \in \mathcal{G}$, where $\mathcal{G} = \mathcal{X} \times [0, T]$ is a compact set in $\mathbb{R}^4$. Given $u \in \mathcal{G}$, D denotes the following 3×3 symmetric matrix-valued function

$$D(u) = \begin{bmatrix} D_{11}(u) & D_{12}(u) & D_{13}(u) \\ D_{12}(u) & D_{22}(u) & D_{23}(u) \\ D_{13}(u) & D_{23}(u) & D_{33}(u) \end{bmatrix}. \quad (1)$$

We require the following assumption:

- (D1) D is symmetric, positive definite, and twice continuously differentiable with the support in $\mathcal{G}$, where it has a simple maximum eigenvalue.

With a slight abuse of notation, D is equivalent to the following 6×1 vector-valued function which consists of six distinct elements of D in (1):

$$D(u) = [D_{11}(u) \ D_{12}(u) \ D_{13}(u) \ D_{22}(u) \ D_{23}(u) \ D_{33}(u)]^T$$

at $u \in \mathcal{G}$. In this paper, we consider a 6×1 vector-valued function D as the longitudinal diffusion tensor. The principal eigenvector of D associated with the largest eigenvalue at $u \in \mathcal{G}$ represents its dominant direction and is denoted as $v(D(\cdot))$. Note that condition (D1) guarantees that $\frac{\partial}{\partial D} v(D(\cdot))$ is a 3×6 well-defined matrix which is differentiable in a neighborhood of D , see Theorem 8.9 in Magnus and Neudecker (2019).

With a slight abuse of notation, we introduce a 3×1 vector-valued function $x : [0, S] \times [0, T] \rightarrow \mathcal{X}$, where $s \in [0, S]$, $S \in \mathbb{R}_+$ is a spatial point (arc length along the curve) and $t \in [0, T]$ is a temporal point that plays a role of a parameter. The longitudinal ensemble of fiber trajectories can be established by solving the following ordinary differential equation (ODE) with the parameter time $t \in [0, T]$:

$$\frac{\partial}{\partial s} x(s, t) = v(D(x(s, t), t)), \quad s \in [0, S], \quad x(0, t) = x_0 \in \mathcal{X}. \quad (2)$$

We require the following assumption:

(D2) For some $T > 0$, $\{(x_0, t), t \in [0, T]\}$ is inside the support of D .

By integration, (2) is equivalent to

$$x(s, t) = x_0 + \int_0^s v(D(x(\xi, t), t)) d\xi. \quad (3)$$

Under (D1) and (D2), $x(s, t)$ exists, is unique, and stays inside $\mathcal{X}$. The longitudinal ensemble of fiber trajectories, $x(s, t)$, is also referred to as an integral curve of the principal eigenvector of the longitudinal diffusion tensor at the given time point. See Coddington and Levinson (1955) for the uniqueness theorem on ODEs with the parameter.

The time rate of change of the longitudinal ensemble of fiber trajectories, $\frac{\partial}{\partial t} x(s, t)$, exists and it can be carried out by Lebesgue's dominated convergence theorem as follows:

$$\begin{aligned} \frac{\partial}{\partial t} x(s, t) &= \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) \frac{\partial}{\partial x} D(x(\xi, t), t) \frac{\partial}{\partial t} x(\xi, t) d\xi \\ &+ \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) \frac{\partial}{\partial t} D(x(\xi, t), t) d\xi, \quad s \in [0, S], t \in [0, T]. \end{aligned} \quad (4)$$

In this paper, we are interested in testing the null hypothesis $H_0 : \frac{\partial}{\partial t} x(s, t) = \mathbf{0}$, where $\mathbf{0}$ denotes a 3×1 vector of 0s, (i.e., the longitudinal ensemble of fiber trajectories is time-invariant) against the alternative hypothesis $H_A : \frac{\partial}{\partial t} x(s, t) \neq \mathbf{0}$ (i.e., the longitudinal ensemble of fiber trajectories is time-varying). When H_0 is true, it implies that the underlying longitudinal diffusion tensor D also does not depend on time t .

D is not directly observed in DTI. One can infer D in (1) from the work of Stejskal and Tanner (1965) whose equation defines

$$y(u; \mathbf{g}) = -b \mathbf{g}^T D(u) \mathbf{g}, \quad u \in \mathcal{G},$$

where $y(u; \mathbf{g}) = \log \left(\frac{A(u; \mathbf{g})}{A_0(u)} \right)$, A is an observed echo amplitude (signal intensity) in the presence of a magnetic field gradient $\mathbf{g} \in \mathbb{R}^3$ at $u \in \mathcal{G}$, and A_0 is a baseline amplitude without any gradient at $u \in \mathcal{G}$. The constant b is referred to as the b -value which includes the gyromagnetic ratio, the gradient duration and separation, and other timing parameters of

the gradient pulse. In DTI, y , b , and g are observed, and hence, one can solve the equation for D .

However, these signal losses are corrupted by noise measurement errors that arise from factors due to head motion and physiological noise during MRI scans. Thus, we consider the following linear model with additive noise errors for the estimation of the diffusion tensor given the set of $N \geq 6$ magnetic field gradients $g_k \in \mathbb{R}^3$, $k = 1, \dots, N$:

$$Y(u) = BD(u) + \Sigma^{1/2}(u)\mathcal{E}, \quad u \in \mathcal{G}, \quad (5)$$

where $Y(u) = [y(u; g_1) \cdots y(u; g_N)]^\top$ is a $N \times 1$ observed vector where the y -values being vertically stacked and B is a $N \times 6$ known tensor of the b -values obtained from $\{g_1, \dots, g_N\}$ during vectorization of D . B is further assumed to be of rank 6. D is a 6×1 unknown vector which represents the longitudinal diffusion tensor at $u \in \mathcal{G}$. Σ is a $N \times N$ unknown symmetric and positive definite tensor. We further assume

(S) Σ is continuous in the support of D .

In the above $\mathcal{E}$ is a $N \times 1$ random vector with a zero mean vector and an identity variance. A similar model in non-longitudinal design is suggested by Carmichael and Sakhanenko (2016).

In the next section, we outline the estimation procedure for the longitudinal diffusion tensor D , its principal eigenvector $v(D(\cdot))$, the longitudinal ensemble of fiber trajectories $x(s, t)$, and its rate of change in time $\frac{\partial}{\partial t}x(s, t)$. Although they are not directly observable in the longitudinal DTI study, they can be estimated and further used in testing our hypothesis.

3 Estimation

3.1 Observations

Longitudinal DWI scans can be viewed as a collection of data on a non-random 4-dimensional space-time grid. When n denotes the number of points in the 4D grid, we have $n = n_x n_t$, where n_x is the number of spatial points in the 3D grid (typically voxels) and n_t is the number of time points for the brain scan. For example, $n_x = 128 \times 128 \times 48 = 786,432$ voxels or $256 \times 256 \times 96 = 6,291,456$ voxels and $n_t = 7$ when the brain is scanned over 7 different time points. Since n is sufficiently large in the longitudinal DTI study, we rather introduce i.i.d. uniformly distributed random variables U_i , $i = 1, \dots, n$, in $\mathcal{G}$ to represent the observed points in the 4D grid. The use of random design simplifies the theoretical derivations since we work with integrals and derivatives of stochastic processes in Sect. 4. One can work with a fixed design, then one would have to consider numerical approximations to integrals and finite differences in place of derivatives similar to those in Sakhanenko et al. (2021).

Based on (5), the observations in the 4D grid are represented by

$$(U_i, Y(U_i)) \text{ with } Y(U_i) = BD(U_i) + \Sigma^{1/2}(U_i)\mathcal{E}_i, \quad i = 1, \dots, n,$$

where $\mathcal{E}_i$, $i = 1, \dots, n$, are i.i.d. $N \times 1$ random vectors having a zero mean vector and an identity variance. $\mathcal{E}_i$, $i = 1, \dots, n$, are independent of U_i , $i = 1, \dots, n$. These observations enable us to estimate both the diffusion tensor D and the noise tensor Σ in 5. In the following sections, we employ a multi-step estimation procedure to achieve the goal of testing the null hypothesis $H_0 : \frac{\partial}{\partial t}x(s, t) = \mathbf{0}$.

3.2 A longitudinal diffusion tensor

First, we compute the ordinary least squares (*ols*) estimate of D which is denoted by $\tilde{D}_{ols}$. It can be decomposed into $\tilde{D}_{ols}(u) = D(u) + \Gamma(u)$ at $u \in \mathcal{G}$, where Γ is a 6×1 random noise vector such that

$$\Gamma(u) = (B^\top B)^{-1} B^\top \Sigma^{1/2}(u) \Xi, \quad u \in \mathcal{G}.$$

We also define $\Gamma_0(u) = (B^\top B)^{-1} B^\top \Sigma^{1/2}(u)$, $u \in \mathcal{G}$. Then we have

$$\mathbb{E}[\Gamma(u)\Gamma^\top(u)] = \Gamma_0(u)\mathbb{E}[\Xi\Xi^\top]\Gamma_0^\top(u) = \Gamma_0(u)\Gamma_0^\top(u), \quad u \in \mathcal{G}.$$

$(B^\top B)^{-1}$ exists since $N \geq 6$ and B has rank 6.

Second, we use the following Nadaraya–Watson type kernel regression estimator (NWE)

$$\hat{D}_n(u) = \frac{1}{nh_n^4} \sum_{i=1}^n \tilde{D}_{ols}(U_i) K\left(\frac{u - U_i}{h_n}\right), \quad u \in \mathcal{G}, \quad (6)$$

where K is a symmetric probability kernel of order 2 and h_n is the bandwidth. One can choose, for example, the standard Gaussian kernel such that

$$K(u) = (2\pi)^{-2} e^{-0.5u^\top u}, \quad u \in \mathbb{R}^4. \quad (7)$$

3.3 A longitudinal ensemble of fiber trajectories and its rate of change over time

A 6×1 vector $\hat{D}_n$ at $u \in \mathcal{G}$ in (6) can be written as a 3×3 symmetric matrix in a similar manner to (1). Third, we take the principal eigenvector of $\hat{D}_n$ associated with the leading eigenvalue at $u \in \mathcal{G}$. Condition (D1), (D2), and the convergence in probability of $\hat{D}_n$ to D in Lemma 1 in Sect. 4 guarantee that $\hat{D}_n$ for large enough n has a simple maximal eigenvalue, $v(\hat{D}_n(\cdot))$ exists, and is unique. Fourth, $x(s, t)$ in (3) can be estimated by the following plug-in estimator:

$$\hat{X}_n(s, t) = x_0 + \int_0^s v(\hat{D}_n(\hat{X}_n(\xi, t), t)) d\xi, \quad s \in [0, S], t \in [0, T], \quad (8)$$

where x_0 is the initial value in $\mathcal{X}$ and $v(\hat{D}_n(\hat{X}_n(s, t), t))$ denotes the principal eigenvector associated with the largest eigenvalue of the NWE in (6) whose $u = (x, t)$ is replaced by $(\hat{X}_n(s, t), t)$ for $s \in [0, S]$, $t \in [0, T]$.

Lastly, the time rate of change $\frac{\partial}{\partial t} x(s, t)$ in (4) can also be estimated by the plug-in estimator. For $s \in [0, S]$, $t \in [0, T]$,

$$\begin{aligned} \frac{\partial}{\partial t} \hat{X}_n(s, t) &= \int_0^s \frac{\partial}{\partial D} v(\hat{D}_n(\hat{X}_n(s, t), t)) \frac{\partial}{\partial x} \hat{D}_n(\hat{X}_n(s, t), t) \frac{\partial}{\partial t} \hat{X}_n(s, t) d\xi \\ &\quad + \int_0^s \frac{\partial}{\partial D} v(\hat{D}_n(\hat{X}_n(s, t), t)) \frac{\partial}{\partial t} \hat{D}_n(\hat{X}_n(s, t), t) d\xi. \end{aligned} \quad (9)$$

Lemmas in Sect. 4 also show the convergence in probability of $\hat{X}_n(s, t)$ to $x(s, t)$ and the convergence in probability of $\frac{\partial}{\partial t} \hat{X}_n(s, t)$ to $\frac{\partial}{\partial t} x(s, t)$ for large enough n . That is, the solution of (8) exists, is unique, and stays in $\mathcal{X}$.

3.4 A longitudinal noise tensor

The estimation of Σ in (5) is done similarly to the estimation of a scale function in a non-parametric heteroscedastic regression model as in chapter 4.3 in Efremovich (2014). First, we take the following residual tensors for $i = 1, \dots, n$:

$$\tilde{\Sigma}_i = [Y(U_i) - B\hat{D}_n(U_i)][Y(U_i) - B\hat{D}_n(U_i)]^\top.$$

Second, we apply the NWE on $\tilde{\Sigma}_i \approx \Sigma(U_i) + \Sigma(U_i)(\mathcal{E}_i \mathcal{E}_i^\top - \mathbb{I})$ where $\mathbb{I}$ denotes an identity matrix. It is denoted as

$$\hat{\Sigma}_n(u) = \frac{1}{nh_n^4} \sum_{i=1}^n \tilde{\Sigma}_i K\left(\frac{u - U_i}{h_n}\right), \quad u \in \mathcal{G}, \quad (10)$$

where K is the same kernel as in (7). Under condition (S), $\hat{\Sigma}_n$ is a consistent estimator of Σ provided that $h_n \rightarrow 0$ and $nh_n^4 \rightarrow \infty$ as $n \rightarrow \infty$. See Efremovich (2014) for reference.

4 Main theorems

Suppose that $nh_n^7 \rightarrow \beta$ as $n \rightarrow \infty$, where $\beta > 0$ is a known fixed number. Then the following lemmas are satisfied.

Lemma 1 Under (D1) and (D2), we have

$$\sup_{u \in \mathcal{G}} |\hat{D}_n(u) - D(u)| \rightarrow 0$$

in probability as $n \rightarrow \infty$.

Lemma 2 Under (D1), (D2), and (S), we have

$$\sup_{s \in [0, S], t \in [0, T]} |\hat{X}_n(s, t) - x(s, t)| \rightarrow 0$$

in probability as $n \rightarrow \infty$.

Lemma 3 Under (D1) and (D2), we have

$$\sup_{u \in \mathcal{G}} \left| \frac{\partial}{\partial x} \hat{D}_n(u) - \frac{\partial}{\partial x} D(u) \right| \rightarrow 0 \text{ and } \sup_{u \in \mathcal{G}} \left| \frac{\partial}{\partial t} \hat{D}_n(u) - \frac{\partial}{\partial t} D(u) \right| \rightarrow 0$$

in probability as $n \rightarrow \infty$.

Lemma 4 Under (D1), (D2), and (S), we have

$$\sup_{s \in [0, S], t \in [0, T]} \left| \frac{\partial}{\partial t} \hat{X}_n(s, t) - \frac{\partial}{\partial t} x(s, t) \right| \rightarrow 0$$

in probability as $n \rightarrow \infty$.

The proofs of Lemmas 1, 2, 3, and 4 are given in the supplementary material.

In this section, Theorem 1 is provided to construct the asymptotic confidence ellipsoids for the longitudinal ensemble of fiber trajectories. Theorem 2 presents the test on whether the longitudinal ensemble of fiber trajectories is time-invariant. Throughout the following Theorems, we use a 3×3 matrix-valued Green's function G with the initial value $G(\xi, \xi, t) =$

$\mathbb{I}$, where $\mathbb{I}$ denotes an identity matrix. For $\xi \in [0, s]$, $s \in [0, S]$ at the parameter $t \in [0, T]$, the Green's function is defined as

$$G(s, \xi, t) = \mathbb{I} + \int_{\xi}^s \frac{\partial}{\partial D} v(D(x(\tau, t), t)) \frac{\partial}{\partial x} D(x(\tau, t), t) G(\tau, \xi, t) d\tau,$$

where G is continuous in (s, ξ) satisfying a Lipschitz condition with respect to $s \in [0, S]$. See Coddington and Levinson (1955) for the existence and use of the Green's function. Based on the Green's function, we use a 3×6 matrix-valued function

$$g(s, \xi, t) = I_{[0 \leq \xi \leq s]} G(s, \xi, t) \frac{\partial}{\partial D} v(D(x(\xi, t), t)),$$

where I denotes an indicator function (i.e., 1 if $0 \leq \xi \leq s$ and 0 otherwise) in the following Theorems.

Theorems 1 and 2 are provided with the proofs of mean and covariance functions. The rest of the proof of Theorems 1 and 2 is given in the supplementary material, which is based on the functional central limit theorem in terms of the convergence of finite dimensional distributions using Lyapunov's condition and the stochastic equicontinuity. See Vaart and Wellner (1996) and Billingsley (1999) for the functional central limit theorem details.

Theorem 1 Assume (D1), (D2), and (S) hold. Suppose that $nh_n^7 \rightarrow \beta$ as $n \rightarrow \infty$, where $\beta > 0$ is a known fixed number. For $s \in [0, S]$, $t \in [0, T]$, the sequence of stochastic processes

$$\sqrt{nh_n^3} (\hat{X}_n(s, t) - x(s, t))$$

converges weakly in the space of $\mathbb{R}^3$ -valued continuous functions on $[0, S]$ given $t \in [0, T]$ to the Gaussian process $\mathcal{GP}(s, t)$, $s \in [0, S]$, $t \in [0, T]$ with the mean function

$$\mu(s, t) = \frac{\sqrt{\beta}}{2} \int_0^S g(s, \xi, t) \int_{\mathbb{R}^4} \left\langle \frac{\partial^2}{\partial u^2} D(x(\xi, t), t) \psi, \psi \right\rangle K(\psi) d\psi d\xi,$$

and the covariance function for all pairs of points $s, s^* \in [0, S]$ at the fixed time $t \in [0, T]$

$$\begin{aligned} C((s, t), (s^*, t)) &= \frac{1}{8\pi\sqrt{\pi}} \int_0^S g(s, \xi, t) [D(x(\xi, t), t) D^\top(x(\xi, t), t) \\ &\quad + \Gamma_0(x(\xi, t), t) \Gamma_0^\top(x(\xi, t), t)] g^\top(s^*, \xi, t) d\xi. \end{aligned}$$

Proof Let us begin with

$$\begin{aligned} y(s, t) &= \hat{X}_n(s, t) - x(s, t) \\ &= \int_0^s [v(\hat{D}_n(\hat{X}_n(\xi, t), t)) - v(D(x(\xi, t), t))] d\xi \\ &= \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) [\hat{D}_n(x(\xi, t), t) - D(x(\xi, t), t)] d\xi \\ &\quad + \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) \frac{\partial}{\partial x} D(x(\xi, t), t) y(\xi, t) d\xi + r(s, t), \end{aligned}$$

where the remainder $r(s, t)$ is

$$\begin{aligned} r(s, t) = & \int_0^s [v(\hat{D}_n(\hat{X}_n(\xi, t), t)) - v(\hat{D}_n(x(\xi, t), t))] d\xi \\ & - \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) \frac{\partial}{\partial x} D(x(\xi, t), t) y(\xi, t) d\xi \\ & + \int_0^s [v(\hat{D}_n(x(\xi, t), t)) - v(D(x(\xi, t), t))] d\xi \\ & - \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) [\hat{D}_n(x(\xi, t), t) - D(x(\xi, t), t)] d\xi. \end{aligned}$$

Let $z(s, t)$ and $\delta(s, t)$ be as follows:

$$\begin{aligned} z(s, t) = & \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) [\hat{D}_n(x(\xi, t), t) - D(x(\xi, t), t)] d\xi \\ & + \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) \frac{\partial}{\partial x} D(x(\xi, t), t) z(\xi, t) d\xi, \\ \delta(s, t) = & \int_0^s \frac{\partial}{\partial D} v(D(x(\xi, t), t)) \frac{\partial}{\partial x} D(x(\xi, t), t) \delta(\xi, t) d\xi + r(s, t). \end{aligned}$$

Then $y(s, t) = z(s, t) + \delta(s, t)$. Using the Green's function G , $z(s, t)$ equals

$$\begin{aligned} z(s, t) = & \int_0^s G(s, \xi, t) \frac{\partial}{\partial D} v(D(x(\xi, t), t)) [\hat{D}_n(x(\xi, t), t) - D(x(\xi, t), t)] d\xi \\ = & \int_0^s g(s, \xi, t) [\hat{D}_n(x(\xi, t), t) - D(x(\xi, t), t)] d\xi, \\ z(0, t) = & 0, \end{aligned}$$

where $g(s, \xi, t) = I_{[0 \leq \xi \leq s]} G(s, \xi, t) \frac{\partial}{\partial D} v(D(x(\xi, t), t))$. Furthermore, $g(s, \xi, t) \in \mathcal{L}$, $s \in [0, S]$, is almost everywhere continuous and bounded on $\mathbb{R}$, where $\mathcal{L}$ is a linear space of functions with the support of g in $[0, S]$.

By letting $\psi = \frac{(x(\xi, t), t) - U}{h_n}$, the mean function of $z(s, t)$ is as follows:

$$\begin{aligned} \mathbb{E}[z(s, t)] = & \int_0^s g(s, \xi, t) \int_{\mathbb{R}^4} [D(x(\xi, t), t) + D((x(\xi, t), t) - h_n \psi) - D(x(\xi, t), t)] \\ & \times K(\psi) d\psi d\xi - \int_0^s g(s, \xi, t) D(x(\xi, t), t) d\xi. \end{aligned}$$

Let $u = (x, t) \in \mathcal{G}$. By Taylor's theorem in a sufficiently small neighborhood of D , we get

$$= \frac{h_n^2}{2} \int_0^s g(s, \xi, t) \int_{\mathbb{R}^4} \left\langle \frac{\partial^2}{\partial u^2} D(x(\xi, t), t) \psi, \psi \right\rangle K(\psi) d\psi d\xi (1 + o_p(1)).$$

The mean function of the limiting Gaussian process $\mathcal{GP}(s, t)$, $s \in [0, S]$, $t \in [0, T]$ can be obtained by $\mu(s, t) = \lim_{n \rightarrow \infty} \sqrt{nh_n^3} \mathbb{E}[z(s, t)]$, where $nh_n^7 \rightarrow \beta$ as $n \rightarrow \infty$.

Similarly, the covariance function of $z(s, t)$ is as follows:

$$\begin{aligned} \text{Cov}[z(s, t), z(s^*, t^*)] &= \frac{1}{nh_n^8} \int_0^S \int_0^S \int_{\mathbb{R}^4} g(s, \xi, t) D(U) K\left(\frac{(x(\xi, t), t) - U}{h_n}\right) \\ &\quad \times K\left(\frac{(x(\eta, t^*), t^*) - U}{h_n}\right) D^\top(U) g^\top(s^*, \eta, t^*) dU d\eta d\xi \\ &\quad + \frac{1}{nh_n^8} \int_0^S \int_0^S \int_{\mathbb{R}^4} g(s, \xi, t) \Gamma_0(U) K\left(\frac{(x(\xi, t), t) - U}{h_n}\right) \\ &\quad \times K\left(\frac{(x(\eta, t^*), t^*) - U}{h_n}\right) \Gamma_0^\top(U) g^\top(s^*, \eta, t^*) dU d\eta d\xi. \end{aligned}$$

By letting $\eta = \xi + \tau h_n$ and $\psi = \frac{(x(\xi, t), t) - U}{h_n}$, we get

$$\begin{aligned} &= \frac{1}{nh_n^3} \int_0^S \int_{-\xi/h_n}^{(S-\xi)/h_n} \int_{\mathbb{R}^4} g(s, \xi, t) D((x(\xi, t), t) - h_n \psi) \\ &\quad \times K(\psi) K\left(\psi + \frac{(x(\xi + \tau h_n, t^*), t^*) - (x(\xi, t), t)}{h_n}\right) \\ &\quad \times D^\top((x(\xi, t), t) - h_n \psi) g^\top(s^*, \xi + \tau h_n, t^*) d\psi d\tau d\xi \\ &\quad + \frac{1}{nh_n^3} \int_0^S \int_{-\xi/h_n}^{(S-\xi)/h_n} \int_{\mathbb{R}^4} g(s, \xi, t) \Gamma_0((x(\xi, t), t) - h_n \psi) \\ &\quad \times K(\psi) K\left(\psi + \frac{(x(\xi + \tau h_n, t^*), t^*) - (x(\xi, t), t)}{h_n}\right) \\ &\quad \times \Gamma_0^\top((x(\xi, t), t) - h_n \psi) g^\top(s^*, \xi + \tau h_n, t^*) d\psi d\tau d\xi. \end{aligned}$$

If $t \neq t^*$, it is close to infinity as $n \rightarrow \infty$ under any density kernel function. For $t = t^*$, we have

$$\frac{(x(\xi + \tau h_n, t), t) - (x(\xi, t), t)}{h_n} \rightarrow (\tau v(D(\xi, t), t), 0) \text{ as } n \rightarrow \infty.$$

Hence, the covariance function of the limiting Gaussian process $\mathcal{GP}(s, t)$, $s \in [0, S]$, $t \in [0, T]$ is

$$\begin{aligned} C((s, t), (s^*, t)) &= \lim_{n \rightarrow \infty} nh_n^3 \text{Cov}[z(s, t), z(s^*, t)] \\ &= \int_0^S \Psi(v(D(x(\xi, t), t))) g(s, \xi, t) [D(x(\xi, t), t) D^\top(x(\xi, t), t) \\ &\quad + \Gamma_0(x(\xi, t), t) \Gamma_0^\top(x(\xi, t), t)] g^\top(s^*, \xi, t) d\xi, \end{aligned}$$

where $\Psi(v(D(\cdot))) = \int_{\mathbb{R}} \int_{\mathbb{R}^4} K(\psi) K(\psi + (\tau v(D(\cdot)), 0)) d\psi d\tau$. For the Gaussian kernel (7), it is easy to show that $\Psi(v(D(\cdot))) = \frac{1}{8\pi\sqrt{\pi}}$.

The rest of the proof can be shown by two parts:

$$(i) \sqrt{nh_n^3} z(s, t) \Rightarrow \mathcal{GP}(s, t), \quad s \in [0, S], t \in [0, T]$$

via the functional central limit theorem.

$$(ii) \sup_{s \in [0, S], t \in [0, T]} |\delta(s, t)| = o_p\left(\frac{1}{\sqrt{nh_n^3}}\right).$$

□

According to Theorem 1, we construct the $100(1 - \alpha)\%$ confidence ellipsoids for the longitudinal ensemble of fiber trajectories $x(s, t)$, $s \in [0, S]$, $t \in [0, T]$, so that the probability of the ellipsoids containing $x(s, t)$ is approximately $1 - \alpha$:

$$P(|C^{-1/2}((s, t), (s, t))[\sqrt{nh_n^3}(\hat{X}_n(s, t) - x(s, t)) - \mu(s, t)]| \leq R_\alpha) \approx 1 - \alpha,$$

where $P(|Z| \leq R_\alpha) = 1 - \alpha$ for a standard normal vector Z in $\mathbb{R}^3$. These confidence ellipsoids represent a coherently oriented fiber population in a specific ROI given any fixed time point at the $100(1 - \alpha)\%$ confidence level.

The main point of this paper is Theorem 2 which establishes the chi-square type test of whether fiber pathways of the population in that ROI are time-invariant or time-varying over the fixed time period. This chi-square type test is based on the Wald's method using the Moore–Penrose pseudoinverse of the covariance matrix of the multivariate normal distribution (Moore 1977). In what follows, $\{s_1, s_2, \dots, s_m\}$ denotes an ordered set of steps on the longitudinal ensemble of fiber trajectories which belong to $[0, S]$ while $\{t_1, t_2, \dots, t_{n_t}\}$ denotes an ordered set of fixed time points (occurring visits for DWI scans) which belong to $[0, T]$.

Theorem 2 Assume (D1), (D2), and (S) hold. Suppose that $nh_n^7 \rightarrow \beta$ as $n \rightarrow \infty$, where $\beta > 0$ is a known fixed number. Let a and b be positive constants. Consider the following testing problem for $0 \leq a < b \leq T$

$$H_0 : \frac{\partial}{\partial t} x(s, t) = \mathbf{0} \text{ versus } H_A : \frac{\partial}{\partial t} x(s, t) \neq \mathbf{0}, \quad s \in [0, S], t \in [a, b].$$

We propose a test statistic under H_0 as

$$\hat{W}_n(s) = \sqrt{nh_n^3} \int_a^b w^\top(t) \frac{\partial}{\partial t} \hat{X}_n(s, t) dt, \quad s \in [0, S],$$

where $w(t)$ is a 3×1 vector-valued and time-dependent weight function. Then the stochastic process $\hat{W}_n(s)$ converges weakly in the space of $\mathbb{R}$ -valued continuous functions on $[0, S]$ to the Gaussian process $\mathcal{GP}(s)$, $s \in [0, S]$ with the mean function

$$\begin{aligned} \mu(s) = & \frac{\sqrt{\beta}}{2} w^\top(b) \int_0^S g(s, \xi, b) \int_{\mathbb{R}^4} \langle \frac{\partial^2}{\partial u^2} D(x(\xi, b), b) \psi, \psi \rangle K(\psi) d\psi d\xi \\ & - \frac{\sqrt{\beta}}{2} w^\top(a) \int_0^S g(s, \xi, a) \int_{\mathbb{R}^4} \langle \frac{\partial^2}{\partial u^2} D(x(\xi, a), a) \psi, \psi \rangle K(\psi) d\psi d\xi \\ & - \frac{\sqrt{\beta}}{2} \int_a^b (w^\top(t))' \int_0^S g(s, \xi, t) \int_{\mathbb{R}^4} \langle \frac{\partial^2}{\partial u^2} D(x(\xi, t), t) \psi, \psi \rangle K(\psi) d\psi d\xi dt, \end{aligned}$$

and the covariance function for all pairs of points $s, s^* \in [0, S]$

$$\begin{aligned} C(s, s^*) = & \frac{1}{8\pi\sqrt{\pi}} w^\top(b) \int_0^S g(s, \xi, b) [D(x(\xi, b), b) D^\top(x(\xi, b), b) \\ & + \Gamma_0(x(\xi, b), b) \Gamma_0^\top(x(\xi, b), b)] g^\top(s^*, \xi, b) d\xi w(b) \\ & + \frac{1}{8\pi\sqrt{\pi}} w^\top(a) \int_0^S g(s, \xi, a) [D(x(\xi, a), a) D^\top(x(\xi, a), a) \\ & + \Gamma_0(x(\xi, a), a) \Gamma_0^\top(x(\xi, a), a)] g^\top(s^*, \xi, a) d\xi w(a). \end{aligned}$$

For finite points $s_1, s_2, \dots, s_m \in [0, S]$, let

$$\hat{W}_0 = \begin{bmatrix} \hat{W}_n(s_1) \\ \hat{W}_n(s_2) \\ \vdots \\ \hat{W}_n(s_m) \end{bmatrix}, \mu_0 = \begin{bmatrix} \mu(s_1) \\ \mu(s_2) \\ \vdots \\ \mu(s_m) \end{bmatrix}, C_0 = \begin{bmatrix} C(s_1, s_1) & \dots & C(s_1, s_m) \\ C(s_2, s_1) & \dots & C(s_2, s_m) \\ \vdots & \dots & \vdots \\ C(s_m, s_1) & \dots & C(s_m, s_m) \end{bmatrix}.$$

Then it is equivalent to $\hat{W}_0 \Rightarrow \text{MVN}(\mu_0, C_0)$ as $n \rightarrow \infty$. Suppose the rank of the covariance matrix C_0 is $\tilde{m} \leq m$. The Wald test of level α rejects H_0 if and only if

$$[\hat{W}_0 - \mu_0]^\top C_0^+ [\hat{W}_0 - \mu_0] > \chi_{\alpha, df=\tilde{m}}^2,$$

where A^+ denotes the Moore–Penrose pseudoinverse of A and $\chi_{\alpha, df=\tilde{m}}^2$ is the critical value of the limiting chi-square distribution with $\tilde{m}$ degrees of freedom at the significance level α .

Proof Under $H_0 : \frac{\partial}{\partial t} x(s, t) = \mathbf{0}$, $s \in [0, S]$, $t \in [a, b]$, we have

$$\int_a^b w^\top(t) \frac{\partial}{\partial t} \hat{X}_n(s, t) dt = \int_a^b w^\top(t) \left[\frac{\partial}{\partial t} \hat{X}_n(s, t) - \frac{\partial}{\partial t} x(s, t) \right] dt,$$

using the method of integration by parts,

$$\begin{aligned} &= w^\top(b) [\hat{X}_n(s, b) - x(s, b)] - w^\top(a) [\hat{X}_n(s, a) - x(s, a)] \\ &\quad - \int_a^b (w^\top(t))' \hat{X}_n(s, t) dt + \int_a^b (w^\top(t))' x(s, t) dt \\ &= z_0(s) + \delta_0(s), \end{aligned}$$

where

$$\begin{aligned} z_0(s) &= w^\top(b) z(s, b) - w^\top(a) z(s, a) - \int_a^b (w^\top(t))' z(s, t) dt, \\ \delta_0(s) &= w^\top(b) \delta(s, b) - w^\top(a) \delta(s, a) - \int_a^b (w^\top(t))' \delta(s, t) dt + r_0(s), \\ r_0(s) &= w^\top(b) r(s, b) - w^\top(a) r(s, a) - \int_a^b (w^\top(t))' r(s, t) dt, \end{aligned}$$

and $z(s, t)$, $\delta(s, t)$, $r(s, t)$ are defined as in the proof of Theorem 1.

The mean function of $z_0(s)$ is $\mathbb{E}[z_0(s)] = w^\top(b) \mathbb{E}[z(s, b)] - w^\top(a) \mathbb{E}[z(s, a)] - \mathbb{E}[\int_a^b (w^\top(t))' z(s, t) dt]$. In a similar manner to the proof of Theorem 1, the mean function of the limiting Gaussian process $\mathcal{GP}(s)$, $s \in [0, S]$, can be easily derived by $\mu(s) = \lim_{n \rightarrow \infty} \sqrt{nh_n^3} \mathbb{E}[z_0(s)]$, where $nh_n^7 \rightarrow \beta$ as $n \rightarrow \infty$.

For the covariance function of $z_0(s)$, we get

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} nh_n^3 w^\top(b) \text{Cov}[z(s, b), z(s^*, b)] w(b) \\
 &= \frac{1}{8\pi\sqrt{\pi}} w^\top(b) \int_0^S g(s, \xi, b) [D(x(\xi, b), b) D^\top(x(\xi, b), b) \\
 &+ \Gamma_0(x(\xi, b), b) \Gamma_0^\top(x(\xi, b), b)] g^\top(s^*, \xi, b) d\xi w(b), \\
 & \lim_{n \rightarrow \infty} nh_n^3 w^\top(a) \text{Cov}[z(s, a), z(s^*, a)] w(a) \\
 &= \frac{1}{8\pi\sqrt{\pi}} w^\top(a) \int_0^S g(s, \xi, a) [D(x(\xi, a), a) D^\top(x(\xi, a), a) \\
 &+ \Gamma_0(x(\xi, a), a) \Gamma_0^\top(x(\xi, a), a)] g^\top(s^*, \xi, a) d\xi w(a), \\
 & \lim_{n \rightarrow \infty} nh_n^3 w^\top(b) \text{Cov}[z(s, b), z(s^*, a)] w(a) = 0, \\
 & \lim_{n \rightarrow \infty} nh_n^3 w^\top(b) \text{Cov}[z(s, b), \int_a^b (w^\top(t))' z(s^*, t) dt] = 0, \\
 & \lim_{n \rightarrow \infty} nh_n^3 w^\top(a) \text{Cov}[z(s, a), \int_a^b (w^\top(t))' z(s^*, t) dt] = 0.
 \end{aligned}$$

Since $\text{Cov}[\int_a^b (w^\top(t))' z(s, t) dt, \int_a^b (w^\top(t))' z(s^*, t) dt] = O(1/nh_n^2)$, we have

$$\lim_{n \rightarrow \infty} nh_n^3 \text{Cov}[\int_a^b (w^\top(t))' z(s, t) dt, \int_a^b (w^\top(t))' z(s^*, t) dt] = 0.$$

Thus, the covariance function of the limiting Gaussian process $\mathcal{GP}(s)$, $s \in [0, S]$ is stated as in Theorem 2.

Similarly to Theorem 1, the rest of the proof is complete by two parts:

$$\text{(i)} \quad \sqrt{nh_n^3} z_0(s) \Rightarrow \mathcal{GP}(s), \quad s \in [0, S]$$

via the functional central limit theorem.

$$\text{(ii)} \quad \sup_{s \in [0, S]} |\delta_0(s)| = o_p\left(\frac{1}{\sqrt{nh_n^3}}\right).$$

□

We remark that the rank of the covariance matrix C_0 can be determined by the use of the truncated singular value decomposition (TSVD). By the TSVD method, we can construct $C_{0,tsvd} = U_{\tilde{m}} \Lambda_{\tilde{m}} V_{\tilde{m}}^\top$, where $U_{\tilde{m}}$, $V_{\tilde{m}}$ denote $m \times \tilde{m}$ unitary matrices, and $\Lambda_{\tilde{m}}$ denotes a $\tilde{m} \times \tilde{m}$ diagonal matrix corresponding to the $\tilde{m}$ leading singular values ($\tilde{m} \leq m$). Then C_0 is replaced by $C_{0,tsvd}$ in Theorem 2 and the rank of $C_{0,tsvd}$ is $\tilde{m}$.

5 Numerical implementation

In Theorems 1 and 2, all ODEs are approximated via Euler's method with a step size of $\Delta > 0$ that satisfies $s_k = s_{k-1} + \Delta$, $k = 1, \dots, m$, with boundary values $s_0 = 0$ and $s_m = S \in \mathbb{R}_+$, where the number of steps is m . Δ is chosen to be sufficiently small so that the local error is proportional to Δ^2 . We use i.i.d. uniformly distributed U_i , $i = 1, \dots, n$, in

$[0, 1]^4$. We set equally spaced time points in $[0, 1]$, that is, $t_j = j/n_t$, $j = 1, \dots, n_t$, where n_t is the number of time points.

Based on Theorem 8.9 in Magnus and Neudecker (2019), partial derivatives of the p th component of the principal eigenvector associated with the largest eigenvalue with respect to D evaluated at $\hat{D}_n$ can be obtained by

$$\begin{aligned} \frac{\partial}{\partial D} v_p \left(\hat{D}_n \left(\hat{X}_n(s_k, t_j), t_j \right) \right) &= (1 - 0.5I_{[r=w]}) \\ &\times [Z^+ \left(\hat{D}_n \left(\hat{X}_n(s_k, t_j), t_j \right) \right)_{pr} v_w \left(\hat{D}_n \left(\hat{X}_n(s_k, t_j), t_j \right) \right) \\ &+ Z^+ \left(\hat{D}_n \left(\hat{X}_n(s_k, t_j), t_j \right) \right)_{pw} v_r \left(\hat{D}_n \left(\hat{X}_n(s_k, t_j), t_j \right) \right)], \end{aligned}$$

where $Z(\hat{D}_n)_{pr} = (\lambda(\hat{D}_n)\mathbb{I} - \hat{D}_n)_{pr}$ for $p, r, w = 1, 2, 3$ and $k = 0, \dots, m$. Recall that I is an indicator function, $\mathbb{I}$ is an identity matrix, and Z^+ is the Moore–Penrose pseudoinverse of Z .

5.1 Theorem 1

Given a fixed time point $t_j = j/n_t$, $j = 1, \dots, n_t$, we set initial values $\hat{X}_n(s_0, t_j) = x_0$, $\hat{\mu}(s_0, t_j) = \mathbf{0}$, and $\hat{C}((s_0, t_j), (s_0, t_j)) = \mathbf{O}$, where $\mathbf{0}$ and $\mathbf{O}$ denote a 3×1 vector of 0s and a 3×3 matrix of 0s, respectively. For $j = 1, \dots, n_t$ and $k = 1, \dots, m$, we approximate

$$\hat{X}_n(s_k, t_j) \approx \hat{X}_n(s_{k-1}, t_j) + \Delta v \left(\hat{D}_n \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \right),$$

and

$$\begin{aligned} \hat{\mu}(s_k, t_j) &\approx \hat{\mu}(s_{k-1}, t_j) + \Delta \frac{\partial}{\partial D} v(\hat{D}_n(\hat{X}_n(s_{k-1}, t_j), t_j)) \\ &\times \frac{\partial}{\partial \mathbf{x}} \hat{D}_n(\hat{X}_n(s_{k-1}, t_j), t_j) \hat{\mu}(s_{k-1}, t_j) \\ &+ \Delta \frac{\partial}{\partial D} v(\hat{D}_n(\hat{X}_n(s_{k-1}, t_j), t_j)) \\ &\times \left[\sum_{i=1}^3 \frac{\partial^2}{\partial x_i^2} \hat{D}_n(\hat{X}_n(s_{k-1}, t_j), t_j) + \frac{\partial^2}{\partial t^2} \hat{D}_n(\hat{X}_n(s_{k-1}, t_j), t_j) \right], \\ \hat{\mu}(s_k, t_j) &\leftarrow \frac{\sqrt{\beta}}{2} \hat{\mu}(s_k, t_j). \end{aligned}$$

For $j = 1, \dots, n_t$ and $k = 1, \dots, m$, we approximate

$$\hat{I}_0 \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \leftarrow (B^\top B)^{-1} B^\top \hat{\Sigma}_n^{1/2} \left(\hat{X}_n(s_{k-1}, t_j), t_j \right).$$

Then we can approximate the covariance function of Theorem 1 when $s = s^*$ as follows for $j = 1, \dots, n_t$ and $k = 1, \dots, m$:

$$\begin{aligned} \hat{C}((s_k, t_j), (s_k, t_j)) &\approx \hat{C}((s_{k-1}, t_j), (s_{k-1}, t_j)) + \Delta \frac{\partial}{\partial D} v \left(\hat{D}_n \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \right) \\ &\quad \times \frac{\partial}{\partial x} \hat{D}_n \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \hat{C}((s_{k-1}, t_j), (s_{k-1}, t_j)) \\ &\quad + \Delta \hat{C}((s_{k-1}, t_j), (s_{k-1}, t_j)) \\ &\quad \times \left[\frac{\partial}{\partial D} v \left(\hat{D}_n \left(\hat{X}_n(s_{l-1}, t_j), t_j \right) \right) \frac{\partial}{\partial x} \hat{D}_n \left(\hat{X}_n(s_{l-1}, t_j), t_j \right) \right]^\top \\ &\quad + \Delta \frac{\partial}{\partial D} v \left(\hat{D}_n \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \right) \\ &\quad \times \left[\hat{D}_n \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \hat{D}_n^\top \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \right. \\ &\quad \left. + \hat{F}_0 \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \hat{F}_0^\top \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \right] \\ &\quad \times \left[\frac{\partial}{\partial D} v \left(\hat{D}_n \left(\hat{X}_n(s_{k-1}, t_j), t_j \right) \right) \right]^\top, \\ \hat{C}((s_k, t_j), (s_k, t_j)) &\leftarrow \frac{1}{8\pi\sqrt{\pi}} \hat{C}((s_k, t_j), (s_k, t_j)). \end{aligned}$$

Let $\hat{C}((s_0, t_j), (s_k, t_j)) = \hat{C}((s_k, t_j), (s_0, t_j)) = \mathbf{0}$ for $k = 1, \dots, m$ and $j = 1, \dots, n_t$. The covariance function of Theorem 1 when $s \neq s^*$ can be approximated by

$$\begin{aligned} \hat{C}((s_k, t_j), (s_{l+1}, t_j)) &\approx \hat{C}((s_k, t_j), (s_l, t_j)) + \Delta \hat{C}((s_k, t_j), (s_l, t_j)) \\ &\quad \times \left[\frac{\partial}{\partial D} v \left(\hat{D}_n \left(\hat{X}_n(s_l, t_j), t_j \right) \right) \frac{\partial}{\partial x} \hat{D}_n \left(\hat{X}_n(s_l, t_j), t_j \right) \right]^\top, \\ \hat{C}((s_k, t_j), (s_{l+1}, t_j)) &\leftarrow \frac{1}{8\pi\sqrt{\pi}} \hat{C}((s_k, t_j), (s_{l+1}, t_j)), \\ \hat{C}((s_{l+1}, t_j), (s_k, t_j)) &\leftarrow \hat{C}((s_k, t_j), (s_{l+1}, t_j)) \end{aligned}$$

for $k = 1, \dots, m, l = k, \dots, m$, and $j = 1, \dots, n_t$.

5.2 Theorem 2

Let $a = i/n_t$ and $b = j/n_t$ where $i, j = 1, \dots, n_t$ and $i < j$. Theorem 2 uses the results of Theorem 1. For $k = 1, \dots, m$, we get

$$\begin{aligned} \hat{W}_n(s_k) &= \sqrt{nh_n^3} \left[w^\top(b) \hat{X}_n(s_k, b) - w^\top(a) \hat{X}_n(s_k, a) - \int_a^b (w^\top(t))' \hat{X}_n(s_k, t) dt \right] \\ \hat{\mu}(s_k) &= w^\top(b) \hat{\mu}(s_k, b) - w^\top(a) \hat{\mu}(s_k, a) - \int_a^b (w^\top(t))' \hat{\mu}(s_k, t) dt, \end{aligned}$$

whereas

$$\hat{C}(s_k, s_l) = w^\top(b) \hat{C}((s_k, b), (s_l, b)) w(b) + w^\top(a) \hat{C}((s_k, a), (s_l, a)) w(a),$$

for $k, l = 1, \dots, m$. All definite integrals are approximated via Simpson's 1/3 rule. In the next two sections, the simulated and real data analyses were performed using MATLAB R2021a.

6 Monte Carlo simulations

In simulations, the sample size $n = 1,216,000$ ($n_x = 40^3$ and $n_t = 19$) was considered. First, we performed 500 Monte Carlo simulations under the following null hypothesis H_0 : we set the thickness of the fiber bundle as $\varepsilon = 0.05$. $x = [x_1 \ x_2 \ x_3]^\top \in [0, 1]^3$ satisfied $|\sqrt{x_1^2 + x_2^2} - 0.5| < \varepsilon$ and $|x_3 - 0.5| < \varepsilon$. At a fixed time point $t_j = j/19$, $j = 1, \dots, 19$, the diffusion tensor was generated by $D(x, t_j) = V(x)\Lambda V^\top(x)$, where

$$V(x) = \begin{bmatrix} \frac{x_2}{\sqrt{x_1^2 + x_2^2}} & \frac{x_1}{\sqrt{x_1^2 + x_2^2}} & 0 \\ -\frac{x_1}{\sqrt{x_1^2 + x_2^2}} & \frac{x_2}{\sqrt{x_1^2 + x_2^2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Under H_0 , the true longitudinal ensemble of fiber trajectories, $x(s, t_j)$, $j = 1, \dots, 19$, remained semi-circular in the x_1x_2 -plane.

Next, we generated 500 simulations under each of the following four alternative hypotheses H_A : at the former 9 time points, we retained H_0 to generate $x(s, t_j)$, $j = 1, \dots, 9$. At the latter 10 time points, given $\varepsilon = 0.05$, we set $x = [x_1 \ x_2 \ x_3]^\top \in [0, 1]^3$ that satisfied

$$\frac{x_1^2}{(0.5 - \varepsilon)^2} + \frac{x_2^2}{(c - \varepsilon)^2} > 1, \quad \frac{x_1^2}{(0.5 + \varepsilon)^2} + \frac{x_2^2}{(c + \varepsilon)^2} < 1$$

for $c \in [0.55, 0.525, 0.475, 0.45]$ and $|x_3 - 0.5| < \varepsilon$. The diffusion tensor was generated by $D(x, t_j) = V(x, c)\Lambda V^\top(x, c)$, $j = 10, \dots, 19$, where

$$V(x, c) = \begin{bmatrix} \frac{x_2/c}{\sqrt{(x_1/0.5)^2 + (x_2/c)^2}} & \frac{x_1/0.5}{\sqrt{(x_1/0.5)^2 + (x_2/c)^2}} & 0 \\ -\frac{x_1/0.5}{\sqrt{(x_1/0.5)^2 + (x_2/c)^2}} & \frac{x_2/c}{\sqrt{(x_1/0.5)^2 + (x_2/c)^2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

for $c \in [0.55, 0.525, 0.475, 0.45]$. Then the corresponding integral curve $x(s, t_j)$ was in the form of a semi-circle for $j = 1, \dots, 9$ while $x(s, t_j)$ was in the form of a semi-ellipse for $j = 10, \dots, 19$ in the x_1x_2 -plane under H_A . Depending on the value of c under H_A , the true longitudinal ensemble of fiber trajectories $x(s, t_j)$ was either stretched or squeezed along the x_2 direction for $j = 10, \dots, 19$.

Both H_0 and H_A met the following conditions: we used 48 gradient directions. The noise tensor $\Sigma(x, t_j)$, $j = 1, \dots, 19$, was specified by a 48×48 constant matrix where diagonal elements were 1 and all the off diagonal elements were 0.5. B in (5) was generated corresponding to the uniform distribution of gradient directions on a unit sphere in each gradient direction. The bandwidth in (6) and (10) was set as $h = 0.0167$ since the bandwidth of the Gaussian kernel should be approximately within $\varepsilon/3$. All ODEs were approximated with the step size $\Delta = 0.015$ and the number of steps $m = 30$. The estimated integral curves $\hat{X}_n(s, t_1)$ and $\hat{X}_n(s, t_{19})$ were shown in Fig. 1. $\hat{X}_n(s, t_1)$ corresponded to the semi-circle under H_0 , whereas $\hat{X}_n(s, t_{19})$ was seen to be the semi-ellipse whose shape was varied by the value of c .

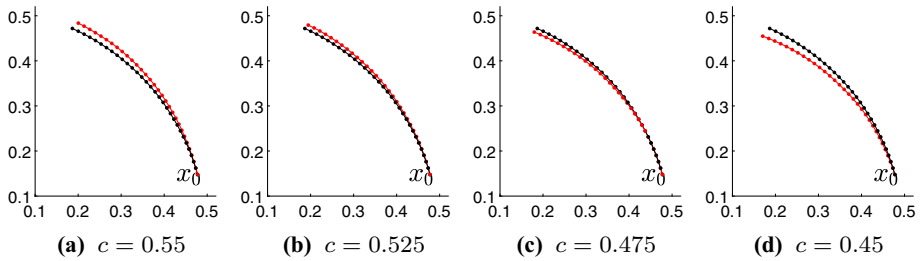


Fig. 1 $\hat{X}_n(s, t_1)$ in black is overlaid with $\hat{X}_n(s, t_{19})$ in red depending on the value of c under H_A . Each step is marked as a dot. All 3D points are projected on the x_1x_2 -plane. (Color figure online)

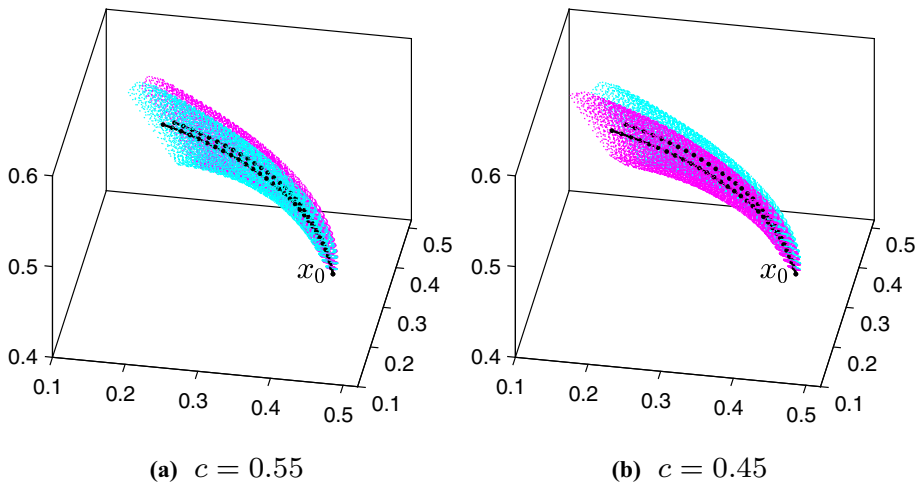


Fig. 2 The 95% estimated confidence ellipsoids for $x(s, t_1)$ in cyan and $x(s, t_{19})$ in magenta are overlaid for the cases of $c = 0.55$ and $c = 0.45$ under H_A . Solid black lines surrounded by the confidence ellipsoids represent $\hat{X}_n(s, t_1)$ and $\hat{X}_n(s, t_{19})$, respectively. Each step is marked as a dot. (Color figure online)

To evaluate the empirical distribution of the test under H_0 in Theorem 2, we considered the following weight functions:

- Linear: $w(t) = [t \ t \ t]^\top$
- Exponential: $w(t) = [e^t \ e^t \ e^t]^\top$
- Constant: $w(t) = [1 \ 1 \ 1]^\top$.

Over 500 simulations under H_0 , we used the TSVD of $\hat{C}_0$ by using the first two singular values since they accounted for more than 98% of the sum of the singular values regardless of the choice of the weight functions. Since the rank of $\hat{C}_{0, \text{tsvd}}$ was 2, the empirical distribution of the corresponding test statistics under H_0 was compared to the chi-square distribution with 2 degrees of freedom. Figure 3 shows that the histogram of these test statistics under H_0 was closely matched with the chi-square distribution with 2 degrees of freedom.

The power of the test was assessed with the theoretical and empirical 5 critical values. The theoretical 5% critical value was obtained from the limiting chi-square distribution with 2 degrees of freedom, whereas the empirical 5% critical value was set to the upper 5th percentile of the 500 simulations under the null hypothesis to ensure a 5% of type I error. In Table 1, we displayed the power based on linear, exponential, constant weight functions. Using the

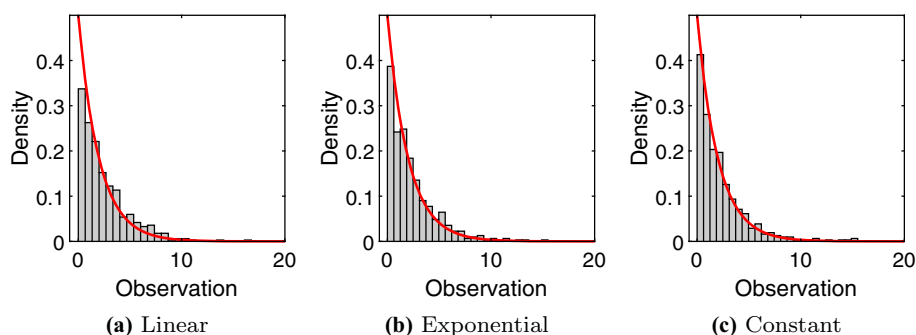


Fig. 3 Histograms of 500 test statistics under H_0 are displayed with different weight functions when $a = t_1$, and $b = t_{19}$. A solid red line is the density function of the chi-square distribution with 2 degrees of freedom. (Color figure online)

Table 1 The power of the test is computed based on both the theoretical 5% critical value ($\chi_{0.05,2}^2 = 5.9915$) and the empirical 5% critical value in parentheses

c	$a = t_1, b = t_{19}$			$a = t_6, b = t_{14}$		
	Linear	Exponential	Constant	Linear	Exponential	Constant
0.55	0.998 (0.998)	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)
0.525	0.832 (0.526)	0.926 (0.676)	0.992 (0.880)	0.936 (0.768)	0.954 (0.772)	0.984 (0.816)
0.475	0.910 (0.688)	0.962 (0.756)	0.998 (0.924)	0.960 (0.828)	0.976 (0.846)	0.986 (0.880)
0.45	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)

linear weight, we weighted less at the time point a and more at the time point b while the time points between were equally weighted since $w'(t) = [1 \ 1 \ 1]^T$. For the exponential weight, $w(t) = w'(t)$, that is, we exponentially weighted over the time period from a through b . For the constant weight, we weighted equally at both end time points a and b , however, middle time points were not considered. The constant weight function yielded the highest power suggesting that it would be the most appropriate when the prior information regarding the weight over time is not available. In Table 1, we also considered two cases of the end time points: $a = t_1, b = t_{19}$ and $a = t_6, b = t_{14}$. When c was either 0.525 or 0.475 (close enough to 0.5 under H_0), the linear weight substantially performed better for the case of $a = t_6, b = t_{14}$ than the case of $a = t_1, b = t_{19}$. This can be explained by the fact that the information at the time point $a = t_1$ was overlooked due to the nature of the linear weight, and hence, the performance of the test was improved when we valued the information at the time point $a = t_6$. The discrepancy between the power based on the empirical critical value and the one based on the theoretical critical value was close to zero when we increased the sample size. The resulting power for the sample size $n = 9,728,000$ ($n_x = 80^3$ and $n_t = 19$) was omitted since it was all one given the aforementioned a, b , weight functions. Since the number of spatial points in DTI, n_x , is much larger than 80^3 , our simulation results imply that the proposed test would detect even a slight variation in the longitudinal ensemble of fiber trajectories with fewer time points.

7 Real data analysis

We analyzed a collection of 19 pre-processed DWI scans on a healthy middle-age male brain from July 2014 to December 2018. During the study period, DWI was performed on a GE 3T Signa HDx MR scanner (GE Healthcare, Waukesha, WI) with an 8-channel head coil. A spin-echo echo-planar imaging sequence (12 min 6 s) was used to acquire the DWI data with the following parameters: 48 contiguous axial slices of 2.4 mm thickness in an interleaved order, $22 \times 22 \text{ cm}^2$ field of view, 2 number of excitations, 13.7 s repetition time, 76.3 ms echo time, 128×128 matrix, 25 diffusion-weighted volumes (one per gradient direction) with $b = 1000 \text{ s/mm}^2$, 1 volume with $b = 0$ and the parallel imaging acceleration factor of 2.

We investigated the posterior part of the corpus callosum (CC) since the CC is the largest among the commissural fiber tracts of a human brain and the posterior of the CC is relatively less affected by the head motion during the MRI scan. A damaged CC can lead to neuronal inter-hemispheric communication issues, and thus severe cognitive dysfunction. For the estimation procedure, we used $\Delta = 0.003$ and $m = 20$ to trace the ensemble of fiber trajectories that curved to the left from $x_0 = [0.5078 \ 0.4219 \ 0.5417]^\top$ and to the right from x_0 . To prevent over- or under-smoothing, we fixed the bandwidth $h_n = 0.01$. The choice of the bandwidth was done in an ad-hoc manner in the estimation step. The effect of the elapsed time between calendar dates is easily handled through reparametrization in this study given that we only observed discrete time points, albeit of a continuous time interval. Therefore, we used the equally spaced time points in $[0, 1]$. Figure 4 shows the longitudinal paths of the fiber population in the posterior region of the CC. We observed that the 95% confidence ellipsoids became enlarged once fibers branched off due to the extremely large estimate of the covariance function. It is the well-known limitation of DTI in complex fiber configurations such as crossing/kissing or branching/merging (Johansen-Berg and Behrens 2014; Jones 2011; Mori and Tournier 2013).

Next, we tested whether there was no time change in the posterior CC at the significance level of 5%. We used $a = t_2$ and $b = t_{18}$ in the posterior CC since the boundary effects were seen when estimating near the end time points such as $a = t_1$ and $b = t_{19}$. Then test statistics were computed by both full and reduced rank approaches corresponding to the considered weight function. The full rank approach used all twenty singular values of $\hat{C}_0$, whereas the reduced rank approach used $\hat{C}_{0,tsvd}$ which was re-constructed by the TSVD method with the first four singular values of $\hat{C}_0$. These four singular values were accounted for more than 96% of the total of singular values of $\hat{C}_0$.

Table 2 Test statistics are shown in two directions of the posterior CC

		Linear	Exponential	Constant
From x_0 to the left	Full rank	3.6395	2.3564	2.6046
	Reduced rank	3.2090	2.0877	2.5191
From x_0 to the right	Full rank	8.3909	3.4841	1.0475
	Reduced rank	7.6037	2.9032	0.7134

In the full rank approach, the critical value is $\chi_{0.05,20}^2 = 31.4104$. In the reduced rank approach, the rank of $\hat{C}_{0,tsvd}$ is 4. The corresponding critical value is $\chi_{0.05,4}^2 = 9.4877$

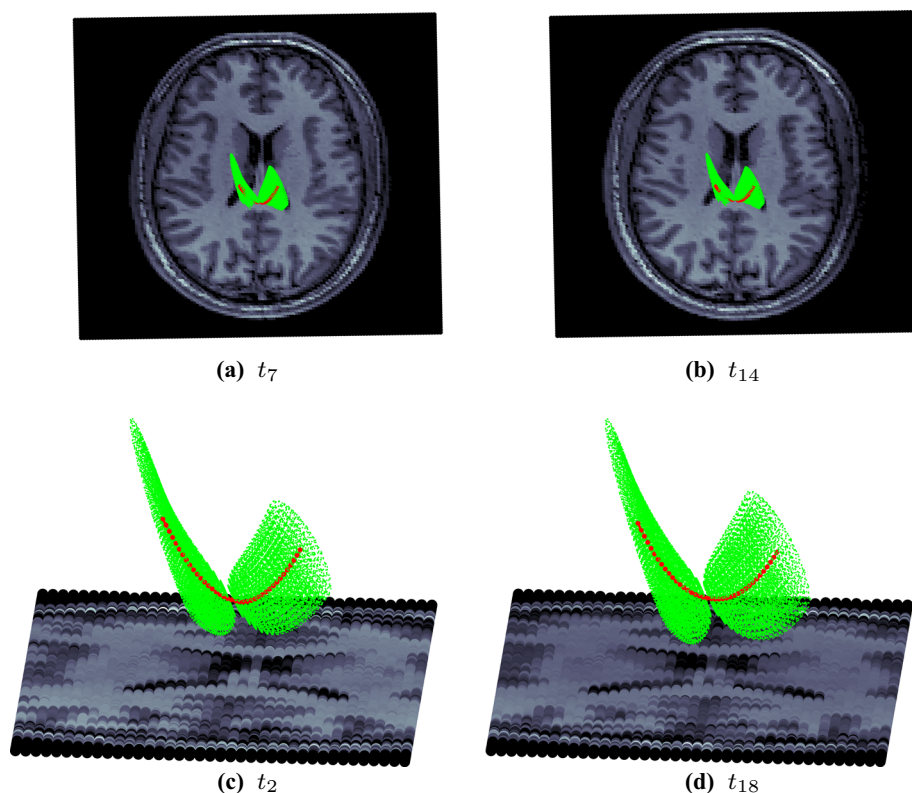


Fig. 4 The 95% estimated confidence ellipsoids in green for the posterior CC is overlaid onto the fractional anisotropy map at a given time point. A solid red line surrounded by the confidence ellipsoids represents the estimated longitudinal ensemble of fiber trajectories at a fixed time point. (Color figure online)

Regardless of the choice of the weight functions and the use of the TSVD method, Table 2 shows no statistically significant evidence to detect the rate of change in the ensemble of fiber trajectories in the posterior CC over the observed period of time. Compared to the simulations in the previous section, we can fairly say that this non-significant result is not due to the lack of power, but it convinces us of no changes in the posterior CC of the healthy brain over the time period.

8 Discussion

In this paper, we are interested in the degree of the time-dependent change in the longitudinal ensemble of fiber trajectories that might be caused by normal aging and neurodegenerative diseases. Unlike the methods of comparing eigenvalues between different time points, the proposed test is directly related to the partial derivative of the longitudinal ensemble of fiber trajectories with respect to time. Furthermore, the proposed approach harnesses all of the information about the time-dependent diffusion tensor in the whole region of the brain. Computationally, it is fairly straightforward and ultimately, it is based on the asymptotic normality of the estimating process. We have demonstrated excellent power and level behavior

of the test on simulated and real data. The proposed test contains tuning parameters such as the constant β , each step size Δ , the number of steps m , and end time points a, b . The effect of tuning parameters on the test should be explored rigorously, which could be a direction for future research.

There are two practical limitations of this test. First, longitudinal DTI studies are often observational studies consisting of very few time points (as low as 2) due to the cost of each MRI scan. In such a case, we can set $[a, b]$ as a short interval and $w(t) = [1 \ 1 \ 1]^T$. Then our test can be based on the finite differences $\hat{X}_n(s, b) - \hat{X}_n(s, a)$, $s \in [0, S]$. One can directly construct tests based on such finite differences, which would be a future direction of our research. Second, the test may not perform well when the data only contains the later stages of the disease. Ideally, one would like to have a baseline information to compare with and a larger number of time points. However, we envision a future when DT-MRI scanning would become a routine prophylactic medical procedure done annually after certain age. Since DTI can potentially be a useful tool to monitor the disease progression over years for high-risk populations of a neurodegenerative disease, such as Alzheimer's disease, our approach can potentially be used to detect the early signs of neuronal fiber damage.

Further investigation of the test should be made on multiple brain locations in many subjects with both healthy and diseased brains. This can be done by the extension to the model in Carmichael and Sakhanenko (2015) as an application of high angular resolution diffusion imaging (HARDI). The fundamental testing problem will remain the same, but the longitudinal HARDI fiber trajectory needs to be defined using the pseudo-eigenvector of a high-order tensor, which would require the development of new mathematical tools. Since the precision of the confidence ellipsoids is higher in HARDI than in DTI, the higher power of the test is expected in HARDI. This would be our future direction of the theoretical research.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11203-022-09268-6>.

Acknowledgements We would like to thank the anonymous reviewers and the associate editor for their thoughtful constructive comments, which led to the improvement of this paper. This research was partially supported by the National Science Foundation grant DMS-1612867, DMS-2111251 and partially supported by Michigan State University High Performance Computing Center through computational resources provided by the Institute for Cyber-Enabled Research.

Data availability The brain image dataset is not publicly available due to confidentiality and the size of the dataset.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Code availability The simulated datasets can be generated by the source code which is available at <https://github.com/jgoo2415/LDTI>.

References

- Billingsley P (1999) Convergence of probability measures, 2nd edn. Wiley, New York
- Carmichael O, Sakhanenko L (2015) Estimation of integral curves from high angular resolution diffusion imaging (HARDI) data. *Linear Algebra Appl* 473:377–403
- Carmichael O, Sakhanenko L (2016) Integral curves from noisy diffusion MRI data with closed-form uncertainty estimates. *Stat Infer Stoch Process* 19:289–319

- Coddington EA, Levinson N (1955) Theory of ordinary differential equations. McGraw-Hill, New York
- Dinkel J., Drier A., Khalilzadeh O., Perlberg V., Czernecki V., Gupta R., Gomas F., Sanchez P., Dormont D., Galanaud D., Stevens R. D., Puybasset L., & for NICER (Neuro Imaging for Coma Emergence and Recovery) Consortium (2014) Long-term white matter changes after severe traumatic brain injury: a 5-year prospective cohort. *Am J Neuroradiol* 35(1):23–29
- Efromovich S (1999) Nonparametric curve estimation: methods, theory and applications. Springer, New York
- Johansen-Berg H, Behrens TEJ (2014) Diffusion MRI: from quantitative measurement to in-vivo, 2nd edn, Neuroanatomy Academic Press
- Jones D (2011) Diffusion MRI: theory, methods, and applications. Oxford University Press, Oxford
- Koltchinskii V, Sakhanenko L, Cai S (2007) Integral curves of noisy vector fields and statistical problems in diffusion tensor imaging: nonparametric kernel estimation and hypotheses testing. *Ann Stat* 35(4):1576–1607
- Kumar M, Kim S, Pickup S, Chen R, Fairless AH, Ittyerah R, Abel T, Brodtkin ES, Poptani H (2012) Longitudinal in-vivo diffusion tensor imaging for assessing brain developmental changes in BALB/cJ mice, a model of reduced sociability relevant to autism. *Brain Res* 1455:56–67
- Magnus JR, Neudecker H (2019) Matrix differential calculus with applications in statistics and econometrics. Wiley, New York
- Moore D (1977) Generalized inverses, Wald's method, and the construction of Chi-squared tests of fit. *J Am Stat Assoc* 72(357):131–137
- Mori S, Tournier JD (2013) Introduction to diffusion tensor imaging: and higher order models. Elsevier, Amsterdam
- Sakhanenko L, DeLaura M, Zhu D (2021) Nonparametric model for a tensor field based on high angular resolution diffusion imaging (HARDI). *Stat Infer Stoch Process* 24:445–476
- Shaffer JJ, Ghayoor A, Long JD, Kim RE, Lourens S, O'Donnell LJ, Westin CF, Rathi Y, Magnotta V, Paulsen JS, Johnson HJ (2017) Longitudinal diffusion changes in prodromal and early HD: evidence of white-matter tract deterioration. *Hum Brain Mapp* 38(3):1460–1477
- Stejskal EO, Tanner JE (1965) Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J Chem Phys* 42:288–292
- Vaart AW, Wellner JA (1996) Weak convergence and empirical processes: with applications to statistics. Springer, New York
- Wu TC, Wilde EA, Bigler ED, Li X, Merkley TL, Yallampalli R, McCauley SR, Schnelle KP, Vasquez AC, Chu Z, Hanten G, Hunter JV, Levin HS (2010) Longitudinal changes in the corpus callosum following pediatric traumatic brain injury. *Dev Neurosci* 32(5–6):361–373
- Zhu DC, Majumdar S, Korolev IO, Berger KL, Bozoki AC (2013) Alzheimer's disease and amnesic mild cognitive impairment weaken connections within the default-mode network: a multi-modal imaging study. *J Alzheimer's Dis* 34(4):969–984