

# Group Comparison of Eigenvalues and Eigenvectors of Diffusion Tensors

Armin SCHWARTZMAN, Robert F. DOUGHERTY, and Jonathan E. TAYLOR

---

Diffusion tensor imaging (DTI) data differ from most medical images in that values at each voxel are not scalars, but  $3 \times 3$  symmetric positive definite matrices called diffusion tensors (DTs). The anatomic characteristics of the tissue at each voxel are reflected by the DT eigenvalues and eigenvectors. In this article we consider the problem of testing whether the means of two groups of DT images are equal at each voxel in terms of the DT's eigenvalues, eigenvectors, or both. Because eigendecompositions are highly nonlinear, existing likelihood ratio statistics (LRTs) for testing differences in eigenvalues or eigenvectors of means of Gaussian symmetric matrices assume an orthogonally invariant covariance structure between the matrix entries. While retaining the form of the LRTs, we derive new approximations to their true distributions when the covariance between the DT entries is arbitrary and possibly different between the two groups. The approximate distributions are those of similar LRT statistics computed on the tangent space to the parameter manifold at the true value of the parameter, but plugging in an estimate for the point of application of the tangent space. The resulting distributions, which are weighted sums of chi-squared distributions, are further approximated by scaled chi-squared distributions by matching the first two moments. For validity of the Gaussian model, the positive definite constraints on the DT are removed via a matrix log transformation, although this is not crucial asymptotically. Voxelwise application of the test statistics leads to a multiple-testing problem, which is solved by false discovery rate inference. The foregoing methods are illustrated in a DTI group comparison of boys versus girls.

**KEY WORDS:** Diffusion tensor imaging; Likelihood ratio test; Manifold-valued data; Multiple testing; Random matrix; Satterthwaite approximation.

---

## 1. INTRODUCTION

Diffusion tensor imaging (DTI) is a novel modality of magnetic resonance imaging that allows in vivo visualization of the internal anatomic structure of the brain's white matter (Basser and Pierpaoli 1996; LeBihan et al. 2001). DTI images are three-dimensional rectangular arrays that contain at every voxel (volume pixel) not a scalar, but a  $3 \times 3$  symmetric positive definite matrix, also called a diffusion tensor (DT). The DT describes the local pattern of water diffusion and can be thought of as the covariance matrix of a three-dimensional Gaussian distribution that models the Brownian motion of the water molecules in the voxel. The displacement of the molecules is not observable; rather, the DT is reconstructed from measurements of the diffusion coefficient in at least six directions in space. The DT serves as a proxy for local anatomic structure. The DT's eigenvalues measure diffusivity and are indicative of the type of tissue and its health, whereas the eigenvectors relate to the spatial orientation of the underlying neural fibers.

A common statistical problem in DTI group studies is to find regions of the brain whose anatomic characteristics differ between two groups of subjects. The analysis typically involves normalizing the images to a common template so that each voxel corresponds to the same anatomic structure in all of the images, and then applying two-sample tests at each voxel. Because of familiarity with univariate statistics, analyses often are restricted to scalar quantities derived from the DT, such as trace and fractional anisotropy (FA), both of which are functions of the DT's eigenvalues and are related to the total diffusivity and the degree of anisotropy within a voxel, respectively. Current

multivariate approaches provide inference for the DT as a single multivariate unit, but not for its eigenvalues and eigenvectors (Basser and Pajevic 2003; Whitcher et al. 2007), or else they focus on the principal diffusion direction (PDD), the eigenvector corresponding to the largest eigenvalue, failing to take into account the role of the eigenvalues in the estimation (Jones et al. 2002; Wu et al. 2004; Schwartzman, Dougherty, and Taylor 2005, 2008).

In this article we introduce a new practical multivariate approach to making inferences about the DT's eigenstructure in group DTI studies. Specifically, given two groups of observed DTs, we consider a test of whether the two group means have the same set of eigenvalues while treating the eigenvectors as nuisance parameters, and a test of whether they have the same frame of eigenvectors while treating the eigenvalues as nuisance parameters. These can be thought of as multivariate extensions of trace/FA tests and PDD tests, respectively. The hypotheses involved are submanifolds of the set of symmetric matrices times itself. A particularly important hypothesis, which plays the role of null and alternative, respectively, in the foregoing tests, is the hypothesis that the means of the two groups have common eigenvalues but possibly different eigenvectors. We note that a related reverse hypothesis appears in the common principal components model of Flury (1984) involving symmetric matrices with common eigenvectors and possibly different eigenvalues, except that those are covariances of Gaussian vectors, rather than means of Gaussian symmetric matrices.

Likelihood ratio test (LRT) statistics for the aforementioned eigenvalue and eigenvector tests have been derived under a Gaussian signal-plus-noise model (Schwartzman, Mascarenhas, and Taylor 2008). However, because eigendecompositions are highly nonlinear functions of the data, those LRTs assume an orthogonally invariant (OI) covariance structure between the

---

Armin Schwartzman is Assistant Professor, Department of Biostatistics, Harvard School of Public Health and Dana-Farber Cancer Institute, Boston, MA 02115 (E-mail: [armins@hsph.harvard.edu](mailto:armins@hsph.harvard.edu)). Robert F. Dougherty is Senior Research Scientist, Department of Psychology (E-mail: [bobd@stanford.edu](mailto:bobd@stanford.edu)) and Jonathan E. Taylor is Associate Professor, Department of Statistics (E-mail: [jonathan.taylor@stanford.edu](mailto:jonathan.taylor@stanford.edu)), Stanford University, Stanford, CA 94305. This work was supported in part by National Institutes of Health grant EY015000 and National Science Foundation grant DMS-0405970.

matrix entries and assume that the covariances in the two groups are the same. The eigenvector test assuming an OI covariance has been useful for identifying white matter tracts affected by necrosis in a brain cancer survivor (Rauschecker et al. 2009). Nevertheless, to allow this test to be used more generally in practical DTI analysis, an arbitrary covariance structure between the tensor elements must be allowed. In this article we retain the form of the LRT statistics derived under the OI assumption but derive new approximations to their distributions when the covariance between the DT entries is arbitrary and possibly different in the two groups. The approximate distributions are those of other LRT statistics obtained when the parameter manifolds are replaced by their tangent spaces at the estimated value of the parameter. These distributions are weighted mixtures of chi-squared distributions, which are then further approximated by scaled chi-squared distributions using the method of moments. The approximations are asymptotically valid as the number of observations per voxel increases.

To ensure validity of the Gaussian model, we assume that the positive definite constraints on the DT data have been removed before analysis by means of a matrix log transformation (Arsigny et al. 2006; Schwartzman 2006; Fletcher and Joshi 2007). The matrix log, computed by taking the log of the eigenvalues and keeping the eigenvectors intact, maps the observed positive definite matrices to the set of symmetric matrices. This is required because the symmetric matrices form a vector space, whereas the positive definite matrices do not. This idea was proposed earlier for modeling covariance matrices (Leonard and Hsu 1992; Chiu, Leonard, and Tsui 1996). Whether the log transform should be applied in practice to DTI data is a subject of current debate (Whitcher et al. 2007). The methods developed in this article remain applicable if the DTs are modeled directly, because the effect of the positive definite constraints disappears asymptotically. In addition, because the matrix log affects only the eigenvalues in a one-to-one fashion, the various hypotheses about eigenvalues and eigenvectors can be equivalently stated in both the original and log domains.

Application of the test statistics at each voxel leads to a multiple-testing problem involving hundreds of thousands of tests. This is addressed using false discovery rate (FDR) inference, now commonly used in the analysis of neuroimaging data (Genovese, Lazar, and Nichols 2002; Logan and Rowe 2004; Schwartzman et al. 2009).

We demonstrate the aforementioned methods in data from an observational study of brain anatomy in children (Dougherty et al. 2007). Here we focus on a cross-sectional analysis whose goal is to find brain regions that differ significantly between boys and girls at age 10. The most important test for this data analysis is the two-sample eigenvector test, which reveals differences in neural fiber orientation between boys and girls mostly in the posterior left hemisphere.

## 2. THEORY

### 2.1 A Gaussian Signal-Plus-Noise Model

Let  $\mathcal{S}_p$  denote the set of  $p \times p$  symmetric matrices ( $p \geq 2$ ). For observed  $\mathbf{Y} \in \mathcal{S}_p$ , let

$$\mathbf{Y} = \mathbf{M} + \mathbf{Z}, \quad (1)$$

where  $\mathbf{M} \in \mathcal{S}_p$  (capital  $\mu$ ) is a mean parameter and  $\mathbf{Z} \in \mathcal{S}_p$  has mean 0. In DTI,  $p = 3$ ,  $\mathbf{Y}$  is the matrix logarithm of the observed diffusion tensor at a particular voxel,  $\mathbf{M}$  represents the population mean at that voxel, and  $\mathbf{Z}$  incorporates both inter-subject variability and measurement noise at that voxel. The effect of measurement noise (Zhu et al. 2007) is not given separate treatment here, because it may be assumed to be negligible compared with the anatomic variability between subjects.

An arbitrary covariance between the entries of  $\mathbf{Y}$  may be specified via the operator

$$\text{vecd}(\mathbf{Y}) = (\text{diag}(\mathbf{Y})', \sqrt{2} \text{offdiag}(\mathbf{Y})')',$$

defined as a column vector of length  $q = p(p + 1)/2$ , where  $\text{diag}(\mathbf{Y})$  is a  $p \times 1$  vector containing the diagonal entries of  $\mathbf{Y}$  and  $\text{offdiag}(\mathbf{Y})$  is a  $(q - p) \times 1$  vector containing the off-diagonal entries of  $\mathbf{Y}$  copied from below the diagonal column-wise (or above the diagonal row-wise). In DTI,  $\text{vecd}(\mathbf{Y}) = (Y_{11}, Y_{22}, Y_{33}, \sqrt{2}Y_{12}, \sqrt{2}Y_{13}, \sqrt{2}Y_{23})'$  and  $q = 6$ . The operator  $\text{vecd}(\cdot)$  was chosen for its convenient property of converting the Frobenius norm of symmetric matrices into the Euclidean norm for vectors, that is,

$$\|\mathbf{Y}\|^2 = \text{tr}(\mathbf{Y}^2) = \text{vecd}(\mathbf{Y})' \text{vecd}(\mathbf{Y}) = \|\text{vecd}(\mathbf{Y})\|^2. \quad (2)$$

If  $\mathbf{Z}$  is nondegenerate Gaussian, then the distribution of  $\mathbf{Y}$ , denoted by  $\mathbf{Y} \sim N_{pp}(\mathbf{M}, \mathbf{\Sigma})$ , may be called symmetric-matrix-variate normal and written as the multivariate normal density

$$\begin{aligned} \mathbf{Y} &\sim N_{pp}(\mathbf{M}, \mathbf{\Sigma}) && \Leftrightarrow \\ \text{vecd}(\mathbf{Y}) &\sim N_q(\text{vecd}(\mathbf{M}), \mathbf{\Sigma}), \end{aligned} \quad (3)$$

where  $\mathbf{\Sigma}$  is a  $q \times q$  positive definite matrix.

Schwartzman, Mascarenhas, and Taylor (2008) derived various LRTs for the mean parameter  $\mathbf{M}$  when  $\mathbf{M}$  is restricted to subsets of  $\mathcal{S}_p$  defined in terms of eigenvalues and eigenvectors of  $\mathbf{M}$ . The LRT statistics derived there assume that the covariance  $\mathbf{\Sigma}$  in (3) is OI. Briefly,  $\mathbf{\Sigma}$  is called OI if the distribution of  $\mathbf{Z} = \mathbf{Y} - \mathbf{M}$  is the same as that of  $\mathbf{QZQ}'$  for any  $\mathbf{Q} \in \mathcal{O}(p)$ , the set of  $p \times p$  orthogonal matrices ( $\mathbf{Q}'\mathbf{Q} = \mathbf{Q}\mathbf{Q}' = \mathbf{I}_p$ ). An OI  $\mathbf{\Sigma}$  can be parameterized by two scalar parameters, a parameter that controls the dependence between the diagonal entries and a global variance parameter (Mallows 1961). A special case for independent entries is the spherical covariance, corresponding to the distribution known in random matrix theory as the Gaussian orthogonal ensemble (GOE) (Mehta 1991).

Although restrictive for data, the OI assumption has the advantage of avoiding the need to vectorize the data matrices, and thus allows one to obtain closed-form solutions to the maximum likelihood (ML) estimates and LRTs for many hypotheses defined in terms of eigenvalues and eigenvectors of  $\mathbf{M}$ . Moreover, Schwartzman, Mascarenhas, and Taylor (2008) showed that in many cases this covariance structure is the only one that allows derivation of closed-form expressions that do not depend on estimates of the covariance parameters. In this article we are interested in applying the LRT statistics derived by Schwartzman, Mascarenhas, and Taylor (2008) when the true covariance  $\mathbf{\Sigma}$  in the data is not necessarily OI.

### 2.2 Two-Sample Tests

Let  $\mathbf{Y}_1, \dots, \mathbf{Y}_{n_1}$  and  $\mathbf{Y}_{n_1+1}, \dots, \mathbf{Y}_n$ ,  $n = n_1 + n_2$ , be two independent iid samples from  $N_{pp}(\mathbf{M}_1, \boldsymbol{\Sigma}_1)$  and  $N_{pp}(\mathbf{M}_2, \boldsymbol{\Sigma}_2)$ , respectively. We consider three tests. The first test, referred to hereinafter as the eigenvalue test, is a test of whether  $\mathbf{M}_1$  and  $\mathbf{M}_2$  have the same eigenvalues, with the eigenvectors unrestricted and treated as nuisance parameters. The test is

$$H_0 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,D} \quad \text{versus} \quad H_A : (\mathbf{M}_1, \mathbf{M}_2) \notin \mathcal{M}_{2,D},$$

where

$$\mathcal{M}_{2,D} = \{(\mathbf{M}_1, \mathbf{M}_2) : \mathbf{M}_1 = \mathbf{U}_1 \mathbf{D} \mathbf{U}_1', \mathbf{M}_2 = \mathbf{U}_2 \mathbf{D} \mathbf{U}_2'\} \quad (4)$$

for unspecified  $\mathbf{D} \in \mathcal{D}_p$ , the set of  $p \times p$  diagonal matrices, and  $\mathbf{U}_1, \mathbf{U}_2 \in \mathcal{O}_p$ . For simplicity,  $\mathbf{D}$  is assumed to have  $p$  distinct eigenvalues.

The second test, referred to hereinafter as the eigenvector test, is a test of whether  $\mathbf{M}_1$  and  $\mathbf{M}_2$  have the same eigenvectors, when the eigenvalues are treated as nuisance parameters and assumed equal between the two populations. The test is

$$H_0 : \mathbf{M}_1 = \mathbf{M}_2 \quad \text{versus} \quad H_A : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,D},$$

where  $\mathcal{M}_{2,D}$  is given by (4).

For comparison, we also consider the test

$$H_0 : \mathbf{M}_1 = \mathbf{M}_2 \quad \text{versus} \quad H_A : \mathbf{M}_1 \neq \mathbf{M}_2,$$

referred to hereinafter as the full matrix test, where no particular attention is paid to the eigenstructure. In this case the  $\mathbf{Y}_i$ 's can be viewed as multivariate samples of dimension  $q$  from model (3). Because  $\boldsymbol{\Sigma}_1$  and  $\boldsymbol{\Sigma}_2$  are not assumed to be equal, this is a multivariate Behrens–Fisher problem. Defining the group averages  $\bar{\mathbf{Y}}_1 = (1/n_1) \sum_{i=1}^{n_1} \mathbf{Y}_i$  and  $\bar{\mathbf{Y}}_2 = (1/n_2) \sum_{i=n_1+1}^n \mathbf{Y}_i$ , Hotelling's  $T^2$  statistic is given by

$$T^2 = \mathbf{d}' \mathbf{S}^{-1} \mathbf{d}, \quad \mathbf{d} = \text{vecd}(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2), \mathbf{S} = \frac{\mathbf{S}_1}{n_1} + \frac{\mathbf{S}_2}{n_2}, \quad (5)$$

where

$$\begin{aligned} \mathbf{S}_1 &= \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} \text{vecd}(\mathbf{Y}_i - \bar{\mathbf{Y}}_1) \text{vecd}(\mathbf{Y}_i - \bar{\mathbf{Y}}_1)', \\ \mathbf{S}_2 &= \frac{1}{n_2 - 1} \sum_{i=n_1+1}^n \text{vecd}(\mathbf{Y}_i - \bar{\mathbf{Y}}_2) \text{vecd}(\mathbf{Y}_i - \bar{\mathbf{Y}}_2)' \end{aligned} \quad (6)$$

are the group sample covariances. The null distribution of (5) may be approximated by redefining

$$T_F = \frac{f - q + 1}{qf} T^2 \underset{H_0}{\sim} F(q, f - q + 1), \quad (7)$$

where  $f$  is the approximate number of degrees of freedom of Yao (1965) given by

$$\frac{1}{f} = \frac{1}{n_1 - 1} \left( \frac{\mathbf{d}' \mathbf{S}^{-1} \mathbf{S}_1 \mathbf{S}^{-1} \mathbf{d}}{n_1 \mathbf{d}' \mathbf{S}^{-1} \mathbf{d}} \right)^2 + \frac{1}{n_2 - 1} \left( \frac{\mathbf{d}' \mathbf{S}^{-1} \mathbf{S}_2 \mathbf{S}^{-1} \mathbf{d}}{n_2 \mathbf{d}' \mathbf{S}^{-1} \mathbf{d}} \right)^2.$$

### 2.3 LRT Statistics

Computing the LRT statistic in both the eigenvalue and eigenvector tests involves estimating the pair  $\mathbf{M} = (\mathbf{M}_1, \mathbf{M}_2)$  under the hypothesis  $\mathbf{M} \in \mathcal{M}_{2,D}$ . The set  $\mathcal{M}_{2,D}$  (4) is a curved submanifold of  $\mathcal{S}_p \times \mathcal{S}_p$ . Assuming that  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$ , Schwartzman, Mascarenhas, and Taylor (2008) showed that the estimation of  $\mathbf{M} \in \mathcal{M}_{2,D}$  is separable from the estimation of  $\boldsymbol{\Sigma}$  if and only if  $\boldsymbol{\Sigma}$  is OI. In that case the ML estimate of  $\mathbf{M}$  is analytically tractable and is the same that would be obtained were  $\boldsymbol{\Sigma}$  assumed to be spherical with variance 1, that is,  $\boldsymbol{\Sigma} = \mathbf{I}_q$ . In other words, the ML estimate is the same as the least squares solution and is found by orthogonal projection.

Specifically, let  $\mathbf{M}_1 = \mathbf{U}_1 \mathbf{D}_1 \mathbf{U}_1'$  and  $\mathbf{M}_2 = \mathbf{U}_2 \mathbf{D}_2 \mathbf{U}_2'$ ,  $\bar{\mathbf{Y}}_1 = \mathbf{V}_1 \boldsymbol{\Lambda}_1 \mathbf{V}_1'$ ,  $\bar{\mathbf{Y}}_2 = \mathbf{V}_2 \boldsymbol{\Lambda}_2 \mathbf{V}_2'$ , and  $\bar{\mathbf{Y}} = \mathbf{V} \boldsymbol{\Lambda} \mathbf{V}'$  be eigendecompositions, all with eigenvalues in decreasing order. Assuming spherical covariances  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \mathbf{I}_q$ , the ML estimate of  $\mathbf{M} = (\mathbf{M}_1, \mathbf{M}_2)$  when  $\mathbf{M} \in \mathcal{M}_{2,D}$  is given by

$$\hat{\mathbf{M}} = (\hat{\mathbf{M}}_1, \hat{\mathbf{M}}_2) = (\hat{\mathbf{U}}_1 \hat{\mathbf{D}} \hat{\mathbf{U}}_1', \hat{\mathbf{U}}_2 \hat{\mathbf{D}} \hat{\mathbf{U}}_2') \in \mathcal{M}_{2,D}, \quad (8)$$

where  $\hat{\mathbf{D}} = \bar{\boldsymbol{\Lambda}} = (n_1 \boldsymbol{\Lambda}_1 + n_2 \boldsymbol{\Lambda}_2)/n$ , and  $\hat{\mathbf{U}}_1$  and  $\hat{\mathbf{U}}_2$  are any matrices of the form  $\hat{\mathbf{U}}_1 = \mathbf{V}_1 \mathbf{Q}_1$  and  $\hat{\mathbf{U}}_2 = \mathbf{V}_2 \mathbf{Q}_2$ , where  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$  are diagonal matrices with diagonal entries equal to  $\pm 1$  (Schwartzman, Mascarenhas, and Taylor 2008, thm. 5.1).

LRT statistics are constructed as follows. In general, suppose that  $\mathcal{M}_0 \subset \mathcal{M}_A \subset \mathcal{S}_p \times \mathcal{S}_p$  are nested hypotheses about the pair  $\mathbf{M} = (\mathbf{M}_1, \mathbf{M}_2)$ . Assuming that  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \mathbf{I}_q$ , the LRT statistic, defined as minus twice the log ratio of the maximized likelihoods under  $\mathcal{M}_A$  and  $\mathcal{M}_0$ , is given by

$$T = \|\bar{\mathbf{Y}} - \hat{\mathbf{M}}_0\|^2 - \|\bar{\mathbf{Y}} - \hat{\mathbf{M}}_A\|^2, \quad (9)$$

where  $\bar{\mathbf{Y}} = (\bar{\mathbf{Y}}_1, \bar{\mathbf{Y}}_2)$ ,  $\hat{\mathbf{M}}_0$  and  $\hat{\mathbf{M}}_A$  are the ML estimates of  $\mathbf{M}$  under  $\mathcal{M}_0$  and  $\mathcal{M}_A$ , and the norm  $\|\cdot\|^2$  in  $\mathcal{S}_p \times \mathcal{S}_p$  is defined in terms of the Frobenius norm (2) by

$$\|\bar{\mathbf{Y}} - \hat{\mathbf{M}}\|^2 = n_1 \|\bar{\mathbf{Y}}_1 - \hat{\mathbf{M}}_1\|^2 + n_2 \|\bar{\mathbf{Y}}_2 - \hat{\mathbf{M}}_2\|^2. \quad (10)$$

For the eigenvalue test, the ML estimate  $\hat{\mathbf{M}}_0$  under  $H_0 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,D}$  is equal to (8). Under  $H_A : (\mathbf{M}_1, \mathbf{M}_2) \notin \mathcal{M}_{2,D}$ , the ML estimate is  $\hat{\mathbf{M}}_A = (\bar{\mathbf{Y}}_1, \bar{\mathbf{Y}}_2)$ . Replacing in (9) gives that, when  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \mathbf{I}_q$ , the LRT statistic is

$$T_D = \frac{n_1 n_2}{n} \|\boldsymbol{\Lambda}_1 - \boldsymbol{\Lambda}_2\|^2 \quad (11)$$

and is asymptotically  $\chi^2(p)$  under  $H_0$  as  $n_1, n_2 \rightarrow \infty$  (Schwartzman, Mascarenhas, and Taylor 2008, cor. 5.1). Not surprisingly, the test statistic measures the distance between the eigenvalue matrices of the two group averages. The number of degrees of freedom corresponds to the fact that  $p$  eigenvalues are being tested.

Similarly, for the eigenvector test, the ML estimate under  $H_0 : \mathbf{M}_1 = \mathbf{M}_2$  is  $\hat{\mathbf{M}}_0 = (\bar{\mathbf{Y}}, \bar{\mathbf{Y}})$ , where  $\bar{\mathbf{Y}} = (n_1 \bar{\mathbf{Y}}_1 + n_2 \bar{\mathbf{Y}}_2)/n$ . Under  $H_A : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,D}$ , the ML estimate  $\hat{\mathbf{M}}_A$  is equal to (8). Replacing in (9) gives that when  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \mathbf{I}_q$ , the LRT statistic is

$$T_U = \frac{2n_1 n_2}{n} [\text{tr}(\boldsymbol{\Lambda}_1 \boldsymbol{\Lambda}_2) - \text{tr}(\bar{\mathbf{Y}}_1 \bar{\mathbf{Y}}_2)] \quad (12)$$

and is asymptotically  $\chi^2(q - p)$  under  $H_0$  as  $n_1, n_2 \rightarrow \infty$  (Schwartzman, Mascarenhas, and Taylor 2008, cor. 5.2). In this

case the number of degrees of freedom corresponds to the fact that the set of eigenvectors being tested has dimension  $q - p$ . The functional form of the test statistic (12) is interesting. When the eigenvectors of  $\bar{\mathbf{Y}}_1$  and  $\bar{\mathbf{Y}}_2$  are equal, the test statistic is equal to 0. As the angles between the eigenvectors of  $\bar{\mathbf{Y}}_1$  and  $\bar{\mathbf{Y}}_2$  increase, with the eigenvalues remaining constant, the inner product  $\text{tr}(\bar{\mathbf{Y}}_1 \bar{\mathbf{Y}}_2)$  decreases, thereby increasing the value of the test statistic.

## 2.4 Distributions Under Arbitrary Covariance

The null distributions of the eigenvalue and eigenvector LRT statistics (11) and (12) given earlier are asymptotic and valid under the assumption that  $\Sigma_1 = \Sigma_2 = \mathbf{I}_q$ . The goal of this section is to derive approximate asymptotic null distributions for those LRT statistics when the true  $\Sigma_1$  and  $\Sigma_2$  are arbitrary. This task presents two difficulties. First, the LRT statistics are highly nonlinear functions of the data, a consequence of the curvature of the parameter sets involved. Second, the eigenvectors of the true covariance matrices may be oblique to the parameter sets.

Our general approach is as follows. For large  $n$ ,  $\bar{\mathbf{Y}} = (\bar{\mathbf{Y}}_1, \bar{\mathbf{Y}}_2)$  is close to  $\underline{\mathbf{M}} = (\mathbf{M}_1, \mathbf{M}_2)$  with high probability, and the effect of the curvature of the parameter manifolds  $\mathcal{M}_0$  and  $\mathcal{M}_A$  near  $\underline{\mathbf{M}}$  becomes negligible. Therefore, the distribution of the LRT statistic (9) is close to the distribution of another LRT statistic  $T^*(\underline{\mathbf{M}})$  computed on the tangent spaces to  $\mathcal{M}_0$  and  $\mathcal{M}_A$  at  $\underline{\mathbf{M}}$ . Denote these tangent spaces by  $\mathcal{T}_0(\underline{\mathbf{M}})$  and  $\mathcal{T}_A(\underline{\mathbf{M}})$ , respectively, and note that  $\mathcal{T}_0(\underline{\mathbf{M}}) \subset \mathcal{T}_A(\underline{\mathbf{M}})$ . Similar to the derivation of (9), the tangent LRT statistic for testing  $\mathcal{T}_0(\underline{\mathbf{M}})$  versus  $\mathcal{T}_A(\underline{\mathbf{M}})$  under the assumption of spherical covariance is

$$T^*(\underline{\mathbf{M}}) = \|\bar{\mathbf{Y}} - \hat{\underline{\mathbf{M}}}_0^*\|^2 - \|\bar{\mathbf{Y}} - \hat{\underline{\mathbf{M}}}_A^*\|^2 = \|\hat{\underline{\mathbf{M}}}_A^* - \hat{\underline{\mathbf{M}}}_0^*\|^2, \quad (13)$$

where  $\hat{\underline{\mathbf{M}}}_0^*$  and  $\hat{\underline{\mathbf{M}}}_A^*$  are the ML estimates of  $\bar{\mathbf{Y}}$  on  $\mathcal{T}_0(\underline{\mathbf{M}})$  and  $\mathcal{T}_A(\underline{\mathbf{M}})$ , respectively, obtained by orthogonal projection (see Figure 1). Define  $\underline{\mathbf{Z}} = \bar{\mathbf{Y}} - \underline{\mathbf{M}}$ . Because the segment  $\hat{\underline{\mathbf{M}}}_A^* - \hat{\underline{\mathbf{M}}}_0^* \subset \mathcal{T}_A(\underline{\mathbf{M}})$  is orthogonal to the segment  $\hat{\underline{\mathbf{M}}}_0^* - \underline{\mathbf{M}} \subset \mathcal{T}_0(\underline{\mathbf{M}}) \subset \mathcal{T}_A(\underline{\mathbf{M}})$ , the tangent LRT statistic (13) also may be written as

$$T^*(\underline{\mathbf{M}}) = \|\text{Proj}_{\mathcal{T}_0^\perp(\underline{\mathbf{M}})}(\text{Proj}_{\mathcal{T}_A(\underline{\mathbf{M}})} \underline{\mathbf{Z}})\|^2, \quad (14)$$

where  $\mathcal{T}_0^\perp(\underline{\mathbf{M}})$  is the subspace orthogonal to  $\mathcal{T}_0(\underline{\mathbf{M}})$  and  $\text{Proj}(\cdot)$  denotes orthogonal projection.

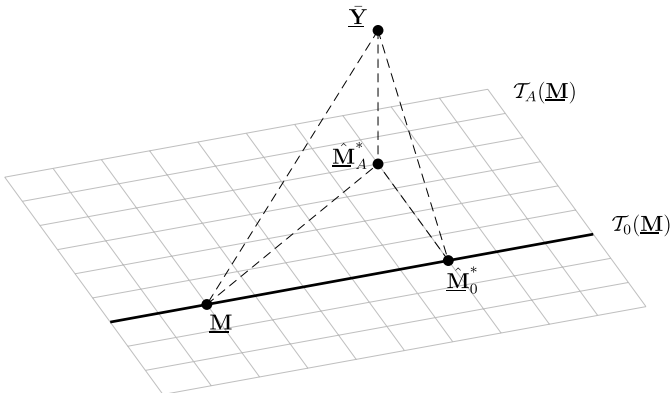


Figure 1. The tangent spaces  $\mathcal{T}_0(\underline{\mathbf{M}}) \subset \mathcal{T}_A(\underline{\mathbf{M}})$  to  $\mathcal{M}_0$  and  $\mathcal{M}_A$  at  $\underline{\mathbf{M}}$  and the corresponding MLEs. The nested underlying manifolds  $\mathcal{M}_0 \subset \mathcal{M}_A$  are not shown.

Given that  $\underline{\mathbf{Z}} = (\mathbf{Z}_1, \mathbf{Z}_2) = (\bar{\mathbf{Y}}_1 - \mathbf{M}_1, \bar{\mathbf{Y}}_2 - \mathbf{M}_2)$ , define the  $2q \times 1$  multivariate normal vector

$$\text{vecd}(\underline{\mathbf{Z}}) = \begin{pmatrix} \text{vecd}(\mathbf{Z}_1) \\ \text{vecd}(\mathbf{Z}_2) \end{pmatrix} \sim N_{2q} \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_1/n_1 & 0 \\ 0 & \Sigma_2/n_2 \end{pmatrix} \right]. \quad (15)$$

Theorems 1 and 2 show that  $T^*(\underline{\mathbf{M}})$  is a quadratic form of  $\text{vecd}(\underline{\mathbf{Z}})$  and give expressions for the specific two-sample tests of interest. For simplified notation, define the elementary matrices  $\mathbf{E}_{ij} = (\mathbf{e}_i \mathbf{e}_j' + \mathbf{e}_j \mathbf{e}_i')/2$ , where  $\mathbf{e}_i$  and  $\mathbf{e}_j$  denote column vectors with a single 1 in positions  $i$  and  $j$ , respectively. Specifically, if  $i = j$ , then  $\mathbf{E}_{ii}$  contains a single 1 in position  $(i, i)$ , and if  $i \neq j$ , then  $\mathbf{E}_{ij} = \mathbf{E}_{ji}$  contains 1/2 in positions  $(i, j)$  and  $(j, i)$ , and 0s elsewhere.

*Theorem 1.* Consider the test  $H_0 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,D}$  versus  $H_A : (\mathbf{M}_1, \mathbf{M}_2) \notin \mathcal{M}_{2,D}$ , where  $\mathcal{M}_{2,D}$  is given by (4). The tangent LRT statistic (14) is

$$T^*(\underline{\mathbf{M}}) = \text{vecd}(\underline{\mathbf{Z}})' \underline{\Omega}(\underline{\mathbf{M}}) \text{vecd}(\underline{\mathbf{Z}}), \quad (16)$$

$$\underline{\Omega}(\underline{\mathbf{M}}) = \frac{n_1 n_2}{n} \sum_{i=1}^p \omega_i \omega_i',$$

where

$$\omega_i = \begin{pmatrix} \text{vecd}(\mathbf{U}_1 \mathbf{E}_{ii} \mathbf{U}_1') \\ -\text{vecd}(\mathbf{U}_2 \mathbf{E}_{ii} \mathbf{U}_2') \end{pmatrix}.$$

*Theorem 2.* Consider the test  $H_0 : \mathbf{M}_1 = \mathbf{M}_2$  versus  $H_A : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,D}$ , where  $\mathcal{M}_{2,D}$  is given by (4). The tangent LRT statistic (14) is

$$T^*(\underline{\mathbf{M}}) = \text{vecd}(\underline{\mathbf{Z}})' \underline{\Omega}(\underline{\mathbf{M}}) \text{vecd}(\underline{\mathbf{Z}}), \quad (17)$$

$$\underline{\Omega}(\underline{\mathbf{M}}) = \frac{n_1 n_2}{n} \sum_{i=1}^p \sum_{j=1}^p \omega_{ij} \omega_{ij}',$$

where

$$\omega_{ij} = \begin{pmatrix} \text{vecd}(\mathbf{E}_{ij}) - \mathbf{J}(\mathbf{U}_1) \mathbf{h}_{ij} \\ -\text{vecd}(\mathbf{E}_{ij}) + \mathbf{J}(\mathbf{U}_2) \mathbf{h}_{ij} \end{pmatrix},$$

$$\mathbf{J}(\mathbf{U}) = (\text{vecd}(\mathbf{U} \mathbf{E}_{11} \mathbf{U}') \quad \text{vecd}(\mathbf{U} \mathbf{E}_{22} \mathbf{U}') \quad \dots \quad \text{vecd}(\mathbf{U} \mathbf{E}_{pp} \mathbf{U}')),$$

$$\mathbf{h}_{ij} = \text{diag}(n_1 \mathbf{U}_2' \mathbf{E}_{ij} \mathbf{U}_2 + n_2 \mathbf{U}_1' \mathbf{E}_{ij} \mathbf{U}_1)/n.$$

In both (16) and (17),  $T^*(\underline{\mathbf{M}})$  has the same form, and its distribution is the same as that of a weighted sum of chi-squared variables, given by Proposition 1. Note that because the distribution of  $\text{vecd}(\underline{\mathbf{Z}})$  does not depend on  $\underline{\mathbf{M}}$ , the dependence on  $\underline{\mathbf{M}}$  is only through  $\underline{\Omega}(\underline{\mathbf{M}})$ . Furthermore,  $\underline{\Omega}(\underline{\mathbf{M}})$  is a function of the eigenvectors of  $\mathbf{M}_1$  and  $\mathbf{M}_2$  only, not their eigenvalues.

*Proposition 1.* Let  $\underline{\Sigma}$  be the covariance matrix of  $\text{vecd}(\underline{\mathbf{Z}})$  given by (15), and let  $k$  be the rank of  $\underline{\Omega}(\underline{\mathbf{M}})$ . Then the distribution of  $T^*(\underline{\mathbf{M}}) = \text{vecd}(\underline{\mathbf{Z}})' \underline{\Omega}(\underline{\mathbf{M}}) \text{vecd}(\underline{\mathbf{Z}})$  is the same as that of  $\mathbf{z}' \underline{\Lambda} \mathbf{z} = \sum_{i=1}^k \lambda_i z_i^2$ , where  $\mathbf{z} \sim N_{2q}(0, \mathbf{I}_{2q})$  and  $\underline{\Lambda}$  is a diagonal matrix containing the  $k$  nonzero eigenvalues of  $\underline{\Sigma}^{1/2} \underline{\Omega}(\underline{\mathbf{M}}) \underline{\Sigma}^{1/2}$ .

The distribution of  $T^*(\mathbf{M})$  can be further approximated as the distribution of a scaled chi-squared variable,  $a\chi_\nu^2$ , matching their first two moments and then solving for  $a, \nu > 0$ , an idea generally known as Welch–Satterthwaite approximation (Stuart and Ord 1994; Kuonen 1999; Casella and Berger 2002). Applying this to the equivalent quadratic form  $\mathbf{z}'\mathbf{A}\mathbf{z}$  in Proposition 1 gives that the first two moments of  $T^*(\mathbf{M})$  are the same as those of a  $a\chi_\nu^2$  variable with

$$a = \frac{\text{tr}(\mathbf{A}^2)}{\text{tr}(\mathbf{A})} = \frac{\text{tr}(\mathbf{\Sigma}\mathbf{\Omega}\mathbf{\Sigma}\mathbf{\Omega})}{\text{tr}(\mathbf{\Sigma}\mathbf{\Omega})}, \quad \nu = \frac{(\text{tr} \mathbf{A})^2}{\text{tr}(\mathbf{A}^2)} = \frac{(\text{tr} \mathbf{\Sigma}\mathbf{\Omega})^2}{\text{tr}(\mathbf{\Sigma}\mathbf{\Omega}\mathbf{\Sigma}\mathbf{\Omega})}.$$

In practice, both  $\mathbf{\Sigma}$  and  $\mathbf{\Omega}(\mathbf{M})$  are unknown. Thus we use a plug-in estimate of  $a$  and  $\nu$  where the required covariances  $\mathbf{\Sigma}_1$  and  $\mathbf{\Sigma}_2$  in (15) are replaced by the sample covariances (6) and  $\mathbf{\Omega}(\mathbf{M})$  is replaced by  $\mathbf{\Omega}(\hat{\mathbf{M}})$ , computed according to (16) and (17) but using the empirical eigenvectors  $\mathbf{V}_1$  and  $\mathbf{V}_2$  of  $\hat{\mathbf{Y}}_1$  and  $\hat{\mathbf{Y}}_2$  instead of the true eigenvectors  $\mathbf{U}_1$  and  $\mathbf{U}_2$  of  $\mathbf{M}_1$  and  $\mathbf{M}_2$ . Given  $T$  from (16) or (17), appropriate  $\hat{a}$  and  $\hat{\nu}$  are found as explained earlier, and the  $p$ -value is computed as  $1 - \hat{F}(T)$ , where  $\hat{F}$  is the cumulative distribution function of  $\hat{a}\chi^2(\hat{\nu})$ .

### 3. NUMERICAL STUDIES

#### 3.1 Distribution

To evaluate the accuracy of the approximate distributions derived earlier, we used the following simulations. Taking  $p = 3$ , two groups of  $n_1 = n_2 = 50$  iid samples were generated according to model (3) with  $\mathbf{M}_0 = \text{Diag}(1, 2, 4)$ , that is, diagonal with diagonal entries 1, 2, and 4, and  $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$  chosen at random as  $\text{Wishart}(I_6, 6)$ . The eigenvalue and eigenvector test statistics (11) and (12) were computed, as were their approximate distributions, as described at the end of Section 2.4. The foregoing procedure was repeated 10,000 times, resulting in 10,000 test statistic values and 10,000 estimated values of  $a$  and  $\nu$  in both cases.

Figure 2 shows the distribution of the  $p$ -values for both test statistics. The proximity of the distribution to the 45-degree line is an indicator of the goodness of the approximation. Similar results were obtained for various instances of the randomly generated  $\mathbf{\Sigma}$ .

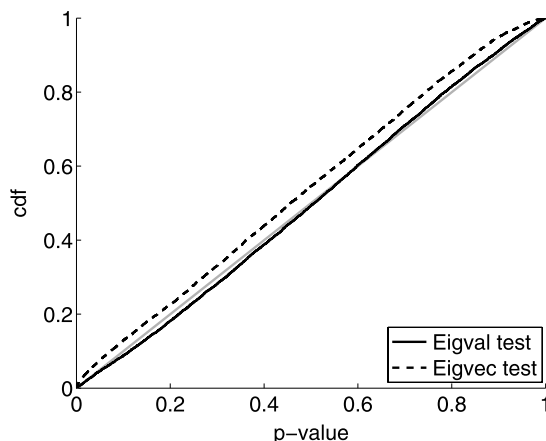


Figure 2. Null distribution of  $p$ -values for the two-sample eigenvalue and eigenvector test statistics, for  $n_1 = 50, n_2 = 50$ .

#### 3.2 Power

The goal of the simulations presented here is to assess the power of the proposed test statistics against various alternative hypotheses. Two groups of  $n_1 = n_2 = 50$  iid samples for  $p = 3$  were generated under the null hypothesis with  $\mathbf{M}_0 = \text{Diag}(1, 2, 4)$  and  $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$  chosen at random as  $\text{Wishart}(I_6, 6)$ , and also under each alternative as described below. Eigenvalue and eigenvector test statistics, along with their approximate tangent space distributions, were computed. The foregoing procedure was repeated 10,000 times, resulting in 10,000 test statistic values and 10,000 estimated  $a$  and  $\nu$  values for each case. The  $p$ -values and statistical power were computed.

The first set of alternatives is defined by changes in the eigenvalues of  $\mathbf{M}_0$  only. These are of the form  $\mathbf{M} = \mathbf{M}_0 + \Delta\mathbf{M}$ , where  $\Delta\mathbf{M}$  is diagonal. Figure 3 summarizes the results for various alternatives in increasing order of difficulty. Alternatives are more difficult to detect when  $\Delta\mathbf{M}$  is smaller, as well as when the eigenvalues of  $\mathbf{M}$  are closer to one another. Interestingly, the eigenvalue test is more powerful than the full matrix test in the harder cases, but not in the easiest case. By design, the eigenvector test has very little power here but is still a little sensitive to changes in eigenvalues, which, with noise, can be mistaken for rotations. It should be noted that the results are strongly dependent on the noise covariance  $\mathbf{\Sigma}$  and the alignment of its eigenvectors with the direction of change of  $\Delta\mathbf{M}$ . Further simulations show that, depending on  $\mathbf{\Sigma}$ , the full matrix test may be more powerful than the eigenvalue test in all four cases.

The second set of alternatives is defined by changes in the eigenvectors. These are of the form  $\mathbf{M} = \mathbf{Q}\mathbf{M}_0\mathbf{Q}'$ , where  $\mathbf{Q} \in \mathcal{O}_3$  is given by Rodrigues’s rotation formula,

$$\mathbf{Q} = \exp \left[ \theta \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \right].$$

Here  $\mathbf{a} = (a_1, a_2, a_3)$  is a unit vector indicating the axis of rotation and  $\theta$  is the rotation angle around that axis. Figure 4 summarizes the results for various alternatives in increasing order of difficulty. In all panels, the eigenvector test is more powerful than the full matrix test. In panel (a), the axis of rotation is oblique to all the eigenvectors of  $\mathbf{M}$ . In panels (b) and (c), rotations are around the axis corresponding to the smallest and largest eigenvalue, respectively. Panel (d) is the same as panel (a) but with noise variance amplified by a factor of 4. By design, the eigenvalue test has no power here—in fact, it is sometimes counterproductive, depending on the particular orientation of the eigenvectors of  $\mathbf{\Sigma}$ . In general, the results depend on  $\mathbf{\Sigma}$ , and further simulations show that the full matrix test sometimes can be more powerful than the eigenvalue test.

### 4. DATA EXAMPLE

#### 4.1 Data Description

Our data set concerns an observational study of brain anatomy in children (Dougherty et al. 2007). A total of 55 children were recruited for the study, and brain scans were taken about once a year. The children were age 8–12 years at the time of the first measurement. For the present study, we extracted a subset of the data consisting of DTI images from only the 10-year-old children. This subset contains 34 brain images

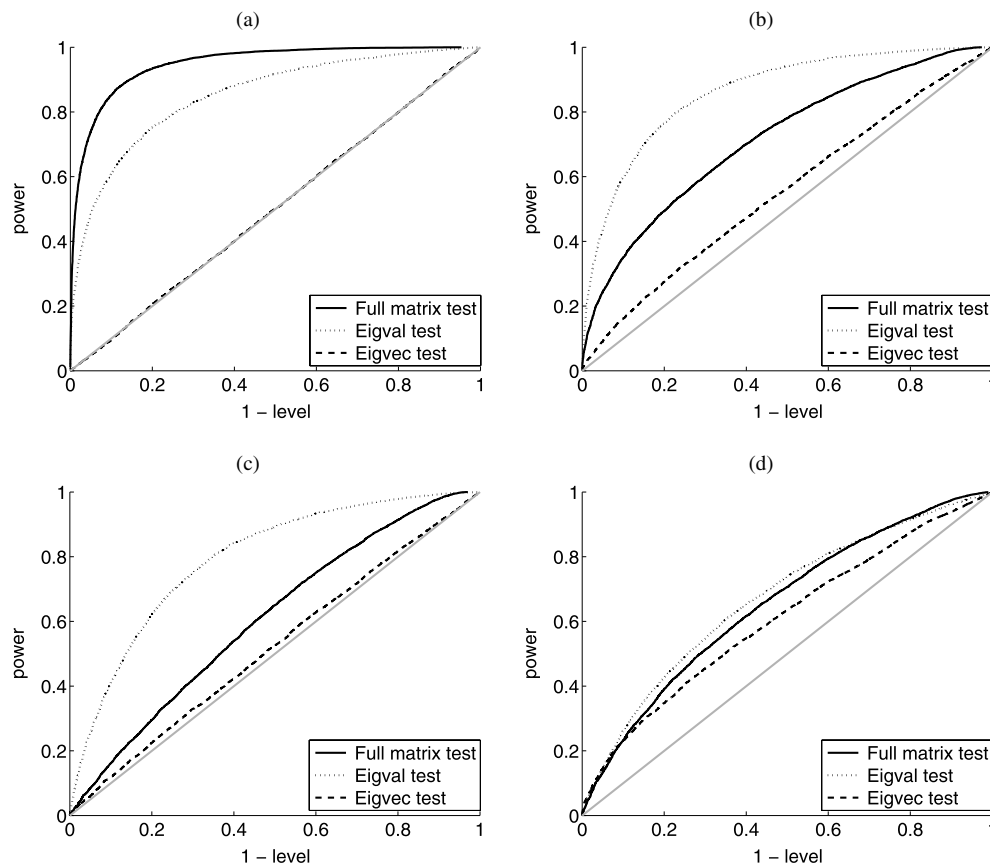


Figure 3. ROC curves for fixed changes in eigenvalues,  $n_1 = 50$ ,  $n_2 = 50$ ,  $\Sigma = \text{Wishart}(\mathbf{I}_6, 6)$ . (a)  $\Delta \mathbf{M} = \text{Diag}(0.2, 0.4, 0.8)$ . (b)  $\Delta \mathbf{M} = \text{Diag}(0.2, 0.4, -0.8)$ . (c)  $\Delta \mathbf{M} = \text{Diag}(0.4, 0.4, -0.4)$ . (d)  $\Delta \mathbf{M} = \text{Diag}(0.4, -0.4, -0.4)$ .

(1 image per child; 12 boys and 22 girls). The goal of the analysis is to find regions of anatomic difference between the boys and the girls.

Voxelwise analysis requires that each voxel correspond to the same brain structure across all subjects. Toward this end, the DTI images were normalized to a common coordinate system based on a custom pediatric template (Dougherty et al. 2005). Normalization resulted in each image being a voxel array of size  $81 \times 106 \times 76$  in a rectangular grid with  $2 \times 2 \times 2$  mm regular spacings. The coordinate axes are defined so that the  $x$ ,  $y$ , and  $z$  axes point to the right, front, and top of the brain, respectively.

Neuroimaging investigators often restrict their analyses to a relevant subset of the brain, called a search region or mask. Here we take a liberal approach and do not restrict the search region to any particular region inside the brain. We computed a whole-brain mask that defines all voxels inside the brain by segmenting each subject's spatially normalized brain and taking the intersection of all such segmentations. The final mask contains  $m = 105,822$  voxels.

#### 4.2 Data Analysis

To eliminate the positive definite constraints before analysis, the eigenvalues of the DT data were log-transformed voxelwise. Next, at each voxel, the two-sample test statistics (7), (11), and (12) were computed, and their approximate null distributions and corresponding  $p$ -values were determined, as described by (7) and discussed in Section 2.4. For our data,  $p = 3$ ,  $q = p(p + 1)/2 = 6$ ,  $n_1 = 12$ , and  $n_2 = 22$ .

Figure 5(a) shows the empirical distribution of the  $p$ -values for the three test types. The eigenvector test produces the highest number of low  $p$ -values, as indicated by the slope of the distribution function at 0. This suggests that the eigenvector test may be more powerful for finding differences between the two groups in this data set.

Figure 6 shows maps of the  $p$ -values at a transverse slice 36 mm above the anterior commissure, an anatomic landmark commonly used for spatial normalization (Talairach convention). Some of the most prominent differences between the two groups are seen in this slice. Regions are highlighted by the three tests to different extents, but the eigenvector test shows the strongest signal.

Determining the significance of the foregoing  $p$ -values is a multiple-testing problem involving  $N = 105,822$  tests. For simplicity, a single threshold was obtained using the classical FDR procedure by Benjamini and Hochberg (1995) at an FDR level of 0.02. Table 1 summarizes the results for the eigenvalue and eigenvector tests. The table reports the  $p$ -value thresholds and the fraction of discoveries  $R/N$ , where  $R$  denotes the number of rejected null hypotheses at that threshold. Here again, the eigenvector test produces many more discoveries.

Figure 7(a) shows the interesting voxels at this FDR level obtained from the eigenvector test at the same slice as in Figure 6. The interesting regions extend to other slices beyond the one shown. For reference, Figure 7(b) shows a map of the total variance at each voxel, that is, the trace of the covariance

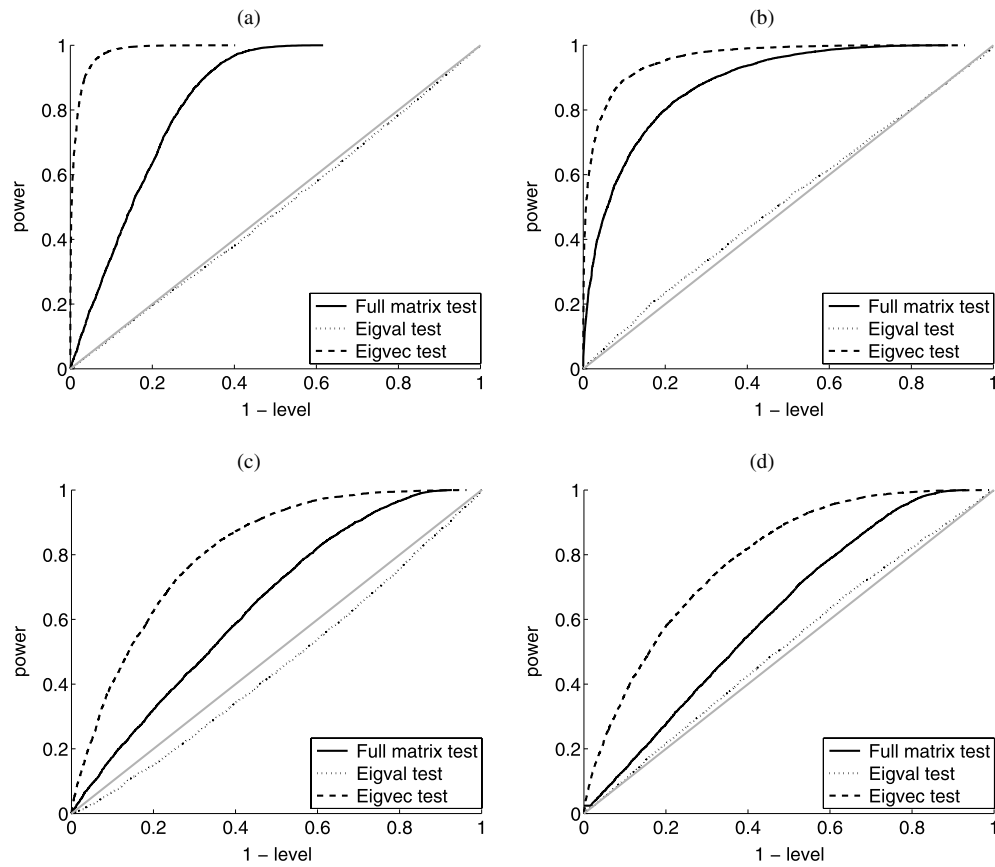


Figure 4. ROC curves for fixed changes in eigenvectors,  $n_1 = 50$ ,  $n_2 = 50$ ,  $\Sigma = \text{Wishart}(\mathbf{I}_6, 6)$ . (a)  $\mathbf{a} = (1, 1, 1)/\sqrt{3}$ ,  $\theta = 0.5$ . (b)  $\mathbf{a} = (1, 0, 0)$ ,  $\theta = 0.5$ . (c)  $\mathbf{a} = (0, 0, 1)$ ,  $\theta = 0.5$ . (d)  $\mathbf{a} = (1, 1, 1)/\sqrt{3}$ ,  $\theta = 0.5$ .

matrix  $S$  given by (5). Compared with the underlying anatomy in panel (a), the lowest variance is seen in the highly coherent neural fibers of the white matter.

Taking advantage of the spatial structure of the data to increase power, we repeated the entire analysis after spatially convolving the log-transformed DT data entry by entry with a uniform box kernel of size  $b \times b \times b$  for  $b = 5$ . Voxels with neighbors outside the mask were eliminated, resulting in a smaller mask with  $N = 73,462$  voxels. Test statis-

tics, approximate distributions, and  $p$ -values were recomputed at each voxel. The new distribution of  $p$ -values is shown in Figure 5(b), and the new FDR thresholds are summarized in the last two columns of Table 1. Both results suggest an increase in power for the eigenvector test, but not necessarily for the eigenvalue test. The  $p$ -value map and corresponding interesting regions are shown in Figure 8 for the same slice as before. Note the difference in texture with Figures 6(c) and 7(a).

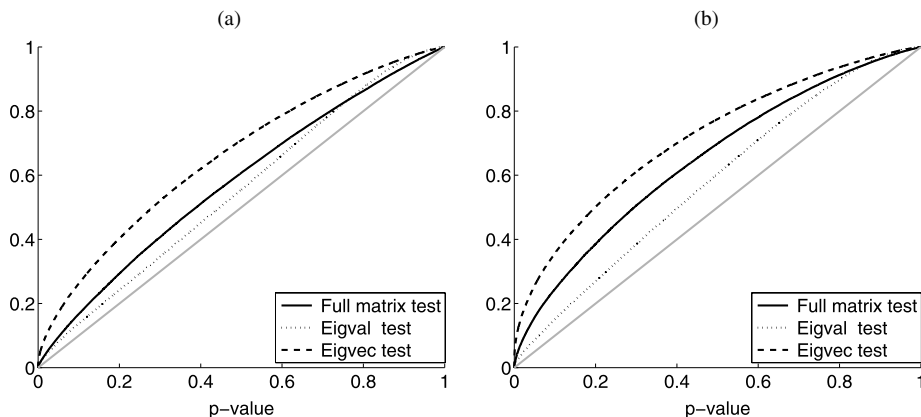


Figure 5. Empirical distribution of  $p$ -values for the three test types: (a) unsmoothed data,  $N = 105,822$  voxels; (b) smoothed data with  $b = 5$ ,  $N = 73,462$  voxels.

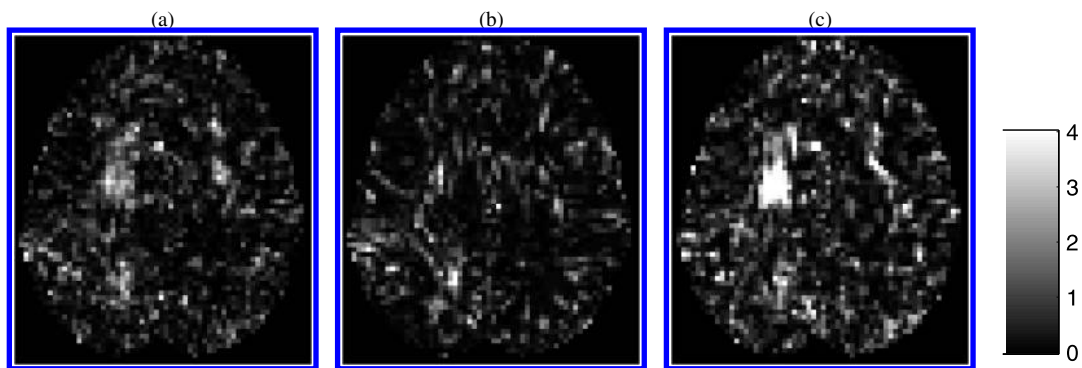


Figure 6.  $p$ -value maps in scale  $-\log_{10}(p)$  at a transverse slice 36 mm above the anterior commissure: (a) full matrix test; (b) eigenvalue test; (c) eigenvector test.

In terms of model checking, a test was applied to the original data to check whether the covariance matrices  $\Sigma_1$  and  $\Sigma_2$  could have been modeled as OI (Schwartzman, Mascarenhas, and Taylor 2008, prop. 3.1). The null hypothesis of OI was rejected at the 0.05 level for 83.8% of the voxels, thereby warranting the need for the approximations developed in Section 2.4. In addition, a Wilks lambda test was applied at each voxel to test the null hypothesis that the group covariance matrices  $\Sigma_1$  and  $\Sigma_2$  are equal. The null hypothesis was rejected at the 0.05 level for 88.6% of voxels, thus warranting the need to assume different covariances between the groups.

For an anatomic interpretation of the results, Figures 9 and 10 show the first and second eigenvectors of the average log-DTs for both groups at the same slice as the previous figures. Because the eigenvector test tests for differences in the full frame of eigenvectors, detected differences may be in the first or second eigenvector (the third is fixed once the first and second are fixed). Many of the regions identified in Figures 7(a) and 8(b) correspond to boundaries between anatomic structures. In the PDD map (first eigenvector), some differences in the crossing patterns can be seen in the left hemisphere between the superior longitudinal fasciculus (SLF) (green) and the corona radiata (blue/purple). Some differences in the crossing patterns also can be seen in this region in the second eigenvector map, which indicate the influence of neural fibers that are not parallel to the PDD. Previous DTI studies of FA have found asymmetries between the two brain hemispheres, particularly in the SLF (Büchel et al. 2004). Our findings encourage further study of the neural connectivity involving this region, possibly by means of DTI tractography.

## 5. DISCUSSION

This article has presented tests for detecting differences in eigenvalues and/or eigenvectors of symmetric matrices between

Table 1. Thresholds and fraction of rejected hypotheses  $R/N$  at FDR level 0.05 for  $N = 105,822$

Test type	Unsmoothed data		Smoothed data, $b = 5$	
	Threshold	$R/N$	Threshold	$R/N$
Eigenvalue	$7.1 \times 10^{-5}$	0.0014	$4.5 \times 10^{-5}$	0.0009
Eigenvector	$1.1 \times 10^{-3}$	0.0225	$4.5 \times 10^{-3}$	0.0895

the means of two groups and has demonstrated their application to comparison of groups of DTs with arbitrary covariance between the DT entries. Of the three tests considered, the eigenvector test was the most powerful for detecting differences between the two groups in the data set analyzed. It is remarkable that the eigenvector test was able to find many regions of difference despite the large multiple-testing problem over 105,822 voxels in the entire brain.

Our methods were applied after a matrix log transformation that maps positive definite matrices to real symmetric matrices. Because the matrix log transformation is bijective and affects only the eigenvalues, the conclusions of the tests are directly interpretable in the original domain. The matrix log transformation is required for validity of the Gaussian model. If desired, the DTs also could be analyzed directly. Asymptotically, the distributions of the group DT averages become more concentrated around their means, and the positive definite constraints have less of an effect. This depends on the Mahalanobis distance from the DTs to the boundaries of the positive definite cone, which in turn is determined by the DT's smallest eigenvalue. Because the trace of the DT is more or less constant over the brain, the matrix log transformation may be more important for highly anisotropic voxels.

The chi-squared null distributions used in this work were derived assuming a Gaussian noise model. A Lilliefors test at each voxel revealed that the distribution across subjects of both the DT entries and the log-DT entries at each voxel was never normal. However, it is easy to see from the derivations that in essence, only the group averages, and not the data, are required to be normally distributed. Asymptotically, this is guaranteed by the central limit theorem and may not be an unreasonable assumption for the moderate sample size in this data set. Another option not explored here is to estimate the null distribution from the data using an empirical null (Efron 2007; Schwartzman 2008).

An advantage of the methodology presented in this article is that test statistics have a closed form and are easy to compute for hundreds of thousands of voxels in the brain. A disadvantage has been the need to find approximations to their null distribution, which we have done with both theoretical and empirical tools. An alternative approach is to directly compute the LRT statistics numerically at each voxel by numerically maximizing the full likelihood. Although much more computationally

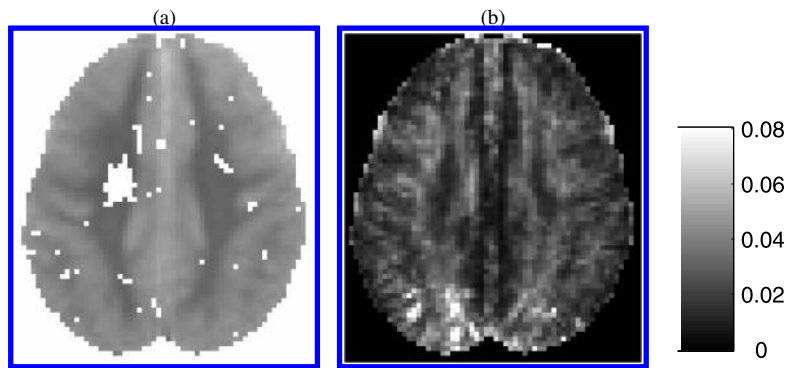


Figure 7. (a) Transverse slice 36 mm above the anterior commissure: Interesting voxels according to eigenvector test at FDR level 0.05. (b) Trace of the empirical covariance  $S = S_1/n_1 + S_2/n_2$ .

costly, it might provide better statistical power, and is worth exploring in future work.

For simultaneous inference across the brain, FDR was chosen for its interpretability and simplicity. As usual in FDR inference, our analysis was based on the marginal distribution of the test statistics. Other approaches, such as random field corrections, common in fMRI studies, make explicit use of the spatial correlation in the data (Worsley et al. 2004). Unlike fMRI data, which have implicit physiological blurring, the tensor measurements are mostly independent by design, with minimal spatial blurring introduced during image acquisition and preprocessing. The distinct patterns of the underlying anatomy, such as fiber tracts, are features of the brain, not artifacts of the measurement. Other recent statistical analyses of DTI data model voxels independently (Zhu et al. 2007). Although some dependency may exist between neighboring voxels, it is of a local nature and qualifies as weak dependence, so FDR control is still guaranteed (Storey, Taylor, and Siegmund 2004).

The spatial structure was exploited to increase the detection power by reanalyzing the data after local spatial smoothing of the log-DT data. This idea was used before in a similar two-group comparison between PDD maps derived from DT data (Schwartzman, Dougherty, and Taylor 2008). That work evaluated a range of sizes,  $b$ , of a uniform box kernel of size  $b \times b \times b$  and found that power increased substantially from  $b = 1$  (no smoothing) to  $b = 3$ , but leveled off at  $b = 7$ . This result prompted our choice of  $b = 5$  as an effective but conservative measure. An important distinction between the two

analyses is that Schwartzman, Dougherty, and Taylor (2008) smoothed the test statistic map rather than the original data. In that work, smoothing the data was problematic, because the data were modeled as Watson vectors. In our case, smoothing the data was appropriate under the Gaussian model, because gaussianity is preserved by linear smoothing.

Although the methods presented in this article have been illustrated in a voxelwise comparison of two groups of subjects, other potential applications in the context of DTI data include analysis of symmetry comparing contralateral brain hemispheres and assessment of error in fiber tracing and spatial registration/normalization.

APPENDIX: PROOFS

Here  $\text{Diag}(\mathbf{A})$  denotes a diagonal matrix with the same diagonal entries as  $\mathbf{A}$ , and  $\text{Offdiag}(\mathbf{A}) = \mathbf{A} - \text{Diag}(\mathbf{A})$  denotes a matrix with the same off-diagonal entries  $\mathbf{A}$  and zero diagonal.

*Lemma A.1.* Denote by  $T_{\mathbf{M}}(\mathcal{M}_{2,D})$  the tangent space to the manifold  $\mathcal{M}_{2,D}$  given by (4) at an arbitrary point  $\mathbf{M} = (\mathbf{U}_1 \mathbf{D} \mathbf{U}'_1, \mathbf{U}_2 \mathbf{D} \mathbf{U}'_2) \in \mathcal{M}_{2,D}$ . Then  $T_{\mathbf{M}}(\mathcal{M}_{2,D})$  coincides with the set of all pairs  $\underline{\mathbf{X}} = (\mathbf{X}_1, \mathbf{X}_2) \in \mathcal{S}_p \times \mathcal{S}_p$  such that

$$\underline{\mathbf{X}} = (\mathbf{U}_1 \mathbf{B}_1 \mathbf{U}'_1 + \mathbf{U}_1 \mathbf{C} \mathbf{U}'_1, \mathbf{U}_2 \mathbf{B}_2 \mathbf{U}'_2 + \mathbf{U}_2 \mathbf{C} \mathbf{U}'_2),$$

where  $\mathbf{B}_1, \mathbf{B}_2 \in \mathcal{S}_p$  have  $\text{diag}(\mathbf{B}_1) = \text{diag}(\mathbf{B}_2) = 0$  and  $\mathbf{C} \in \mathcal{D}_p$ .

*Proof.* Let

$$\underline{\mathbf{M}}(t) = (\mathbf{U}_1 e^{\mathbf{A}_1 t} (\mathbf{D} + \mathbf{C}t) e^{-\mathbf{A}_1 t} \mathbf{U}'_1, \mathbf{U}_2 e^{\mathbf{A}_2 t} (\mathbf{D} + \mathbf{C}t) e^{-\mathbf{A}_2 t} \mathbf{U}'_2)$$

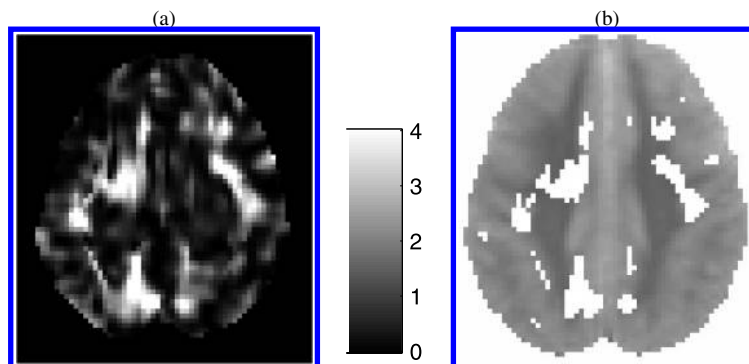


Figure 8. Spatial smoothing of the data with  $b = 5$ , transverse slice 36 mm above the anterior commissure: (a) Map of  $p$ -values in scale  $-\log_{10}(p)$ . (b) Interesting voxels at FDR level 0.05.

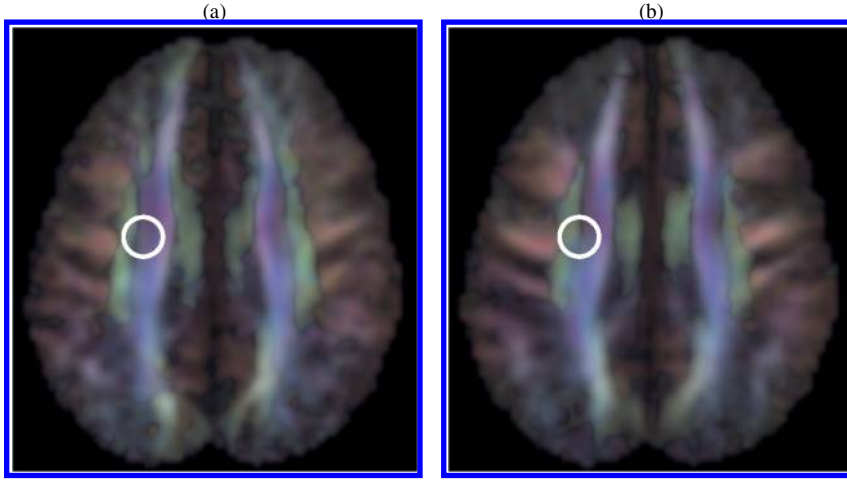


Figure 9. First eigenvector of DT mean for boys (a) and girls (b) at a transverse slice 36 mm above the anterior commissure. Colors indicate coordinate directions: right–left (red) and anterior–posterior (green), and superior–inferior (blue). Mixed colors represent directions that are oblique to the coordinate axes. The intensity is weighted by FA, so brighter regions have higher anisotropy. A color version of this figure is available in the electronic version of this article.

be an arbitrary curve in  $\mathcal{M}_{2,D}$  passing through  $\underline{\mathbf{M}}$  at  $t = 0$ , where  $\mathbf{C} \in \mathcal{D}_p$  and  $\mathbf{A}_1, \mathbf{A}_2 \in \mathcal{A}_p$ , the set of  $p \times p$  antisymmetric matrices. This parameterization relies on the fact that  $e^{A t}$  parameterizes  $\mathcal{O}(p)$  (Edelman, Arias, and Smith 1998; Lang 1999; Moakher 2002). Taking the derivative with respect to  $t$  and evaluating at  $t = 0$  gives that any tangent vector  $\underline{\mathbf{X}} \in \mathcal{T}_{\underline{\mathbf{M}}}(\mathcal{M}_{2,D})$  has the form

$$\underline{\mathbf{X}} = (\mathbf{U}_1(\mathbf{A}_1\mathbf{D} - \mathbf{D}\mathbf{A}_1)\mathbf{U}'_1 + \mathbf{U}_1\mathbf{C}\mathbf{U}'_1, \mathbf{U}_2(\mathbf{A}_2\mathbf{D} - \mathbf{D}\mathbf{A}_2)\mathbf{U}'_2 + \mathbf{U}_2\mathbf{C}\mathbf{U}'_2).$$

Letting  $\mathbf{B}_1 = \mathbf{A}_1\mathbf{D} - \mathbf{D}\mathbf{A}_1$  and  $\mathbf{B}_2 = \mathbf{A}_2\mathbf{D} - \mathbf{D}\mathbf{A}_2$ , it is easy to see that both  $\mathbf{B}_1$  and  $\mathbf{B}_2$  are symmetric matrices with zero diagonal.

Conversely, the  $\underline{\mathbf{X}}$ 's span the tangent space. This is because for every  $\mathbf{B}$  symmetric with zero diagonal, there exists an antisymmetric  $\mathbf{A}$  such that  $\mathbf{B} = \mathbf{A}\mathbf{D} - \mathbf{D}\mathbf{A}$ . To see this, observe that the entry  $(i, j)$ ,  $i \neq j$ , of  $\mathbf{A}\mathbf{D} - \mathbf{D}\mathbf{A}$  is

$$(\mathbf{A}\mathbf{D} - \mathbf{D}\mathbf{A})_{ij} = \mathbf{e}'_i(\mathbf{A}\mathbf{D} - \mathbf{D}\mathbf{A})\mathbf{e}_j = \mathbf{e}'_i\mathbf{A}d_{jj}\mathbf{e}_j - \mathbf{e}'_i d_{ii}\mathbf{A}\mathbf{e}_j = (d_{jj} - d_{ii})a_{ij},$$

where  $\mathbf{e}_i$  denotes a column vector with a single 1 in position  $i$ ,  $d_{ii}$  are the diagonal entries of  $\mathbf{D}$ , and  $a_{ij}$  are the entries of  $\mathbf{A}$ . Therefore, for every  $\mathbf{B}$  with entries  $b_{ij}$ , the corresponding  $\mathbf{A}$  has entries  $a_{ij} = b_{ij}/(d_{jj} - d_{ii})$ ,  $i \neq j$ .

#### Proof of Theorem 1

For this test,  $\mathcal{T}_0(\underline{\mathbf{M}}) = \mathcal{T}_{\underline{\mathbf{M}}}(\mathcal{M}_{2,D})$  and  $\mathcal{T}_A(\underline{\mathbf{M}}) = \mathcal{S}_p \times \mathcal{S}_p$ , so (14) reduces to  $T^*(\underline{\mathbf{M}}) = \|\text{Proj}_{\mathcal{T}_0^\perp(\underline{\mathbf{M}})} \underline{\mathbf{Z}}\|^2$ . By Lemma A.1, the projection of  $\underline{\mathbf{Z}}$  onto  $\mathcal{T}_{\underline{\mathbf{M}}}(\mathcal{M}_{2,D})$  is the minimizer of the square distance

$$\begin{aligned} \|\underline{\mathbf{Z}} - \underline{\mathbf{X}}\|^2 &= \|(\mathbf{Z}_1, \mathbf{Z}_2) - (\mathbf{U}_1\mathbf{B}_1\mathbf{U}'_1 + \mathbf{U}_1\mathbf{C}\mathbf{U}'_1, \mathbf{U}_2\mathbf{B}_2\mathbf{U}'_2 + \mathbf{U}_2\mathbf{C}\mathbf{U}'_2)\|^2 \\ &= n_1\|\mathbf{U}'_1\mathbf{Z}_1\mathbf{U}_1 - \mathbf{B}_1 - \mathbf{C}\|^2 + n_2\|\mathbf{U}'_2\mathbf{Z}_2\mathbf{U}_2 - \mathbf{B}_2 - \mathbf{C}\|^2. \end{aligned}$$

For simplicity, let  $\mathbf{W}_1 = \mathbf{U}'_1\mathbf{Z}_1\mathbf{U}_1$  and  $\mathbf{W}_2 = \mathbf{U}'_2\mathbf{Z}_2\mathbf{U}_2$ . We can write

$$\begin{aligned} \|\underline{\mathbf{Z}} - \underline{\mathbf{X}}\|^2 &= n_1\|\text{Diag}(\mathbf{W}_1 - \mathbf{C})\|^2 + 2n_1\|\text{Offdiag}(\mathbf{W}_1 - \mathbf{B}_1)\|^2 \\ &\quad + n_2\|\text{Diag}(\mathbf{W}_2 - \mathbf{C})\|^2 + 2n_2\|\text{Offdiag}(\mathbf{W}_2 - \mathbf{B}_2)\|^2. \end{aligned}$$

This is minimized over  $\mathbf{B}_1$ ,  $\mathbf{B}_2$ , and  $\mathbf{C}$  for

$$\begin{aligned} \tilde{\mathbf{B}}_1 &= \text{Offdiag}(\mathbf{W}_1), & \tilde{\mathbf{B}}_2 &= \text{Offdiag}(\mathbf{W}_2), \\ \tilde{\mathbf{C}} &= [n_1\text{Diag}(\mathbf{W}_1) + n_2\text{Diag}(\mathbf{W}_2)]/n. \end{aligned} \tag{A.1}$$

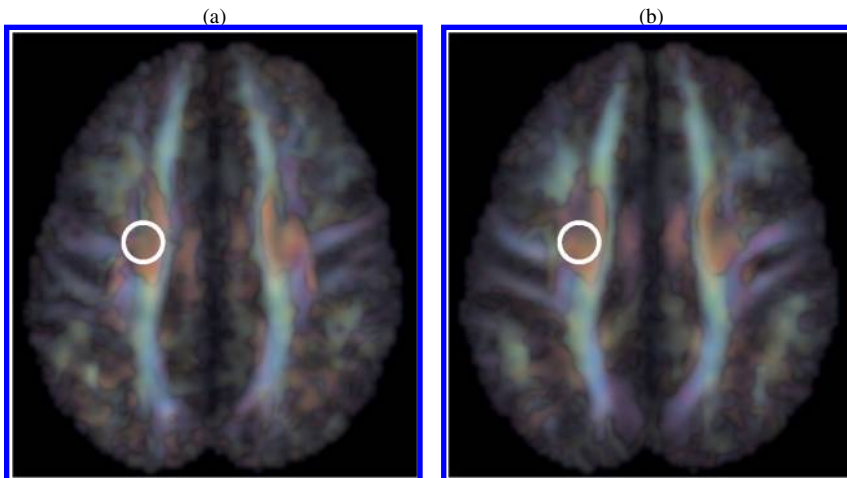


Figure 10. Second eigenvector of DT mean for boys (a) and girls (b) at the same slice and using the same color scheme as Figure 9. A color version of this figure is available in the electronic version of this article.

The norm  $T^*(\mathbf{M})$  of the projection of  $\mathbf{Z}$  onto the space orthogonal to  $\mathcal{T}_{\mathbf{M}}(\mathcal{M}_{2,D})$  is the same as the minimal distance of  $\mathbf{Z}$  to  $\mathcal{T}_{\mathbf{M}}(\mathcal{M}_{2,D})$ . Replacing the optimal  $\tilde{\mathbf{X}}$  with  $\tilde{\mathbf{B}}_1, \tilde{\mathbf{B}}_2$ , and  $\tilde{\mathbf{C}}$  in the expression of the distance gives that  $T^*(\mathbf{M}) = \|\mathbf{Z} - \tilde{\mathbf{X}}\|^2$  is equal to

$$\begin{aligned} T^*(\mathbf{M}) &= \|\mathbf{Z} - \tilde{\mathbf{X}}\|^2 \\ &= n_1 \|\text{Diag}(\mathbf{W}_1 - \tilde{\mathbf{C}})\|^2 + n_2 \|\text{Diag}(\mathbf{W}_2 - \tilde{\mathbf{C}})\|^2 \\ &= \frac{n_1 n_2}{n} \|\text{Diag}(\mathbf{W}_1 - \mathbf{W}_2)\|^2 \\ &= \frac{n_1 n_2}{n} \sum_{i=1}^p \{\text{tr}[\mathbf{e}'_i(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{e}_i]\}^2 \\ &= \frac{n_1 n_2}{n} \sum_{i=1}^p [\text{tr}(\mathbf{Z}_1 \mathbf{U}_1 \mathbf{e}_i \mathbf{e}'_i \mathbf{U}'_1) - \text{tr}(\mathbf{Z}_2 \mathbf{U}_2 \mathbf{e}_i \mathbf{e}'_i \mathbf{U}'_2)]^2 \\ &= \frac{n_1 n_2}{n} \sum_{i=1}^p [\text{vecd}(\mathbf{Z}_1)' \text{vecd}(\mathbf{U}_1 \mathbf{e}_i \mathbf{e}'_i \mathbf{U}'_1) \\ &\quad - \text{vecd}(\mathbf{Z}_2)' \text{vecd}(\mathbf{U}_2 \mathbf{e}_i \mathbf{e}'_i \mathbf{U}'_2)]^2, \end{aligned}$$

which yields the result.

**Proof of Theorem 2**

In the computation of (14) for this test,  $\mathcal{T}_0(\mathbf{M})$  is the tangent space to the set  $\mathcal{M}_{2,M} = \{\mathbf{M} = (\mathbf{M}, \mathbf{M}) : \mathbf{M} \in \mathcal{S}_p\}$ , which can be identified with  $\mathcal{M}_{2,M}$  itself, whereas  $\mathcal{T}_A(\mathbf{M}) = \mathcal{T}_{\mathbf{M}}(\mathcal{M}_{2,D})$ . From the proof of Theorem 1, we have that the projection of  $\mathbf{Z}$  on  $\mathcal{T}_A(\mathbf{M})$  is

$$\text{Proj}_{\mathcal{T}_A(\mathbf{M})} \mathbf{Z} = (\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2) = (\mathbf{U}_1 \tilde{\mathbf{B}}_1 \mathbf{U}'_1 + \mathbf{U}_1 \tilde{\mathbf{C}} \mathbf{U}'_1, \mathbf{U}_2 \tilde{\mathbf{B}}_2 \mathbf{U}'_2 + \mathbf{U}_2 \tilde{\mathbf{C}} \mathbf{U}'_2),$$

where  $\tilde{\mathbf{B}}_1, \tilde{\mathbf{B}}_2$ , and  $\tilde{\mathbf{C}}$  are given by (A.1). By Lemma A.1, the projection of  $\text{Proj}_{\mathcal{T}_A(\mathbf{M})} \mathbf{Z}$  onto  $\mathcal{T}_A(\mathbf{M}) = \mathcal{T}_{\mathbf{M}}(\mathcal{M}_{2,M})$  is the minimizer of the square distance

$$\|(\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2) - (\mathbf{M}, \mathbf{M})\|^2 = n_1 \|\tilde{\mathbf{X}}_1 - \mathbf{M}\|^2 + n_2 \|\tilde{\mathbf{X}}_2 - \mathbf{M}\|^2.$$

This is minimized over  $\mathbf{M} \in \mathcal{S}_p$  for  $\tilde{\mathbf{M}} = (n_1 \tilde{\mathbf{X}}_1 + n_2 \tilde{\mathbf{X}}_2)/n$ . The norm  $T^*(\mathbf{M})$  of the projection of  $\text{Proj}_{\mathcal{T}_A(\mathbf{M})} \mathbf{Z}$  onto the space orthogonal to  $\mathcal{T}_0(\mathbf{M})$  is the same as the minimal distance of  $\text{Proj}_{\mathcal{T}_A(\mathbf{M})} \mathbf{Z}$  to  $\mathcal{T}_0(\mathbf{M})$ . Replacing the optimal  $\tilde{\mathbf{M}}$  in the expression of the distance gives

$$\begin{aligned} T^*(\mathbf{M}) &= \|(\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2) - (\tilde{\mathbf{M}}, \tilde{\mathbf{M}})\|^2 \\ &= n_1 \|\tilde{\mathbf{X}}_1 - \tilde{\mathbf{M}}\|^2 + n_2 \|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{M}}\|^2 = \frac{n_1 n_2}{n} \|\tilde{\mathbf{X}}_1 - \tilde{\mathbf{X}}_2\|^2. \end{aligned}$$

Replacing  $\tilde{\mathbf{X}}_1$  and  $\tilde{\mathbf{X}}_2$  using (A.1), and adding and subtracting  $\mathbf{U}_1 \text{Diag}(\mathbf{W}_1) \mathbf{U}'_1$  and  $\mathbf{U}_2 \text{Diag}(\mathbf{W}_2) \mathbf{U}'_2$ , we get that  $T^*(\mathbf{M})$  can be written as

$$T^*(\mathbf{M}) = \frac{n_1 n_2}{n} \|\mathbf{Z}_1 - \mathbf{H}(\mathbf{W}_1) - \mathbf{Z}_2 + \mathbf{H}(\mathbf{W}_2)\|^2, \tag{A.2}$$

where

$$\mathbf{H}(\mathbf{W}_1) = (n_2 \mathbf{U}_1 \text{Diag}(\mathbf{W}_1) \mathbf{U}'_1 + n_1 \mathbf{U}_2 \text{Diag}(\mathbf{W}_1) \mathbf{U}'_2)/n,$$

$$\mathbf{H}(\mathbf{W}_2) = (n_2 \mathbf{U}_1 \text{Diag}(\mathbf{W}_2) \mathbf{U}'_1 + n_1 \mathbf{U}_2 \text{Diag}(\mathbf{W}_2) \mathbf{U}'_2)/n.$$

The entries of  $\mathbf{H}(\mathbf{W}_1)$  are

$$\begin{aligned} \text{tr}(\mathbf{e}'_i \mathbf{H}(\mathbf{W}_1) \mathbf{e}_i) &= \text{tr}[\text{Diag}(\mathbf{W}_1)(n_2 \mathbf{U}'_1 \mathbf{e}_i \mathbf{e}'_i \mathbf{U}_1 + n_1 \mathbf{U}'_2 \mathbf{e}_i \mathbf{e}'_i \mathbf{U}_2)/n] \\ &= \text{diag}(\mathbf{W}_1)' \mathbf{h}_{ij}, \end{aligned}$$

where  $\mathbf{h}_{ij}$  is the  $p \times 1$  vector,

$$\mathbf{h}_{ij} = \text{diag}(n_1 \mathbf{U}'_2 \mathbf{E}_{ij} \mathbf{U}_2 + n_2 \mathbf{U}'_1 \mathbf{E}_{ij} \mathbf{U}_1)/n.$$

Furthermore, we can write  $\text{diag}(\mathbf{W}_1) = \mathbf{J}(\mathbf{U}_1)' \text{vecd}(\mathbf{Z}_1)$ , where

$$\mathbf{J}(\mathbf{U}) = (\text{vecd}(\mathbf{U}\mathbf{E}_{11}\mathbf{U}') \quad \text{vecd}(\mathbf{U}\mathbf{E}_{22}\mathbf{U}') \quad \dots \quad \text{vecd}(\mathbf{U}\mathbf{E}_{pp}\mathbf{U}')).$$

This can be verified by checking that the entries of the  $p \times 1$  vector  $\text{diag}(\mathbf{W}_1)$  are

$$\begin{aligned} \mathbf{e}'_i \text{Diag}(\mathbf{W}_1) \mathbf{e}_i &= \text{tr}(\mathbf{e}'_i \mathbf{U}_1 \mathbf{Z}_1 \mathbf{U}'_1 \mathbf{e}_i) = \text{tr}(\mathbf{U}'_1 \mathbf{e}_i \mathbf{e}'_i \mathbf{U}_1 \mathbf{Z}_1) \\ &= \text{vecd}(\mathbf{U}'_1 \mathbf{E}_{ii} \mathbf{U}_1)' \text{vecd}(\mathbf{Z}_1). \end{aligned}$$

Similar expressions are obtained for  $\mathbf{H}(\mathbf{W}_2)$ . Going back to (A.2), we have that

$$\begin{aligned} T^*(\mathbf{M}) &= \frac{n_1 n_2}{n} \sum_{i=1}^p \sum_{i=1}^p [\text{tr}(\mathbf{e}'_i \mathbf{Z}_1 \mathbf{e}_i - \mathbf{e}'_i \mathbf{H}(\mathbf{W}_2) \mathbf{e}_i - \mathbf{e}'_i \mathbf{Z}_2 \mathbf{e}_i + \mathbf{e}'_i \mathbf{H}(\mathbf{W}_2) \mathbf{e}_i)]^2 \\ &= \frac{n_1 n_2}{n} \sum_{i=1}^p \sum_{i=1}^p [\text{vecd}(\mathbf{Z}_1)' \text{vecd}(\mathbf{e}_i \mathbf{e}'_i) - \text{vecd}(\mathbf{Z}_1)' \mathbf{J}(\mathbf{U}_1) \mathbf{h}_{ij} \\ &\quad - \text{vecd}(\mathbf{Z}_2)' \text{vecd}(\mathbf{e}_i \mathbf{e}'_i) - \text{vecd}(\mathbf{Z}_2)' \mathbf{J}(\mathbf{U}_2) \mathbf{h}_{ij}]^2, \end{aligned}$$

which gives the result.

**Proof of Proposition 1**

This can be seen as a special case of theorem 3.4.4a of *Mardia, Kent, and Bibby (1979)*, but it also can be proven directly as follows. Let  $\mathbf{z} = \underline{\Sigma}^{-1/2} \text{vecd}(\mathbf{Z}) \sim N_{2q}(0, \mathbf{I}_{2q})$ . We have

$$T^* = \text{vecd}(\mathbf{Z})' \underline{\Omega} \text{vecd}(\mathbf{Z}) = \mathbf{z}' \underline{\Sigma}^{1/2} \underline{\Omega} \underline{\Sigma}^{1/2} \mathbf{z} = (\mathbf{Q}\mathbf{z})' \underline{\Lambda} (\mathbf{Q}\mathbf{z}),$$

where  $\underline{\Lambda}$  and  $\mathbf{Q}$  are the eigenvalue and eigenvector matrices of  $\underline{\Sigma}^{1/2} \underline{\Omega} \underline{\Sigma}^{1/2}$ . The result follows because  $\mathbf{Q}\mathbf{z} \sim N_{2q}(0, \mathbf{I}_{2q})$ .

[Received May 2007. Revised September 2009.]

**REFERENCES**

Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2006), "Log-Euclidean Metrics for Fast and Simple Calculus on Diffusion Tensors," *Magnetic Resonance in Medicine*, 56, 411–421. [589]

Basser, P. J., and Pajevic, S. (2003), "A Normal Distribution for Tensor-Valued Random Variables: Applications to Diffusion Tensor MRI," *IEEE Transactions on Medical Imaging*, 22, 785–794. [588]

Basser, P. J., and Pierpaoli, C. (1996), "Microstructural and Physiological Features of Tissues Elucidated by Quantitative-Diffusion-Tensor MRI," *Journal of Magnetic Resonance, Ser. B*, 111, 209–219. [588]

Benjamini, Y., and Hochberg, Y. (1995), "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society, Ser. B*, 57, 289–300. [593]

Büchel, C., Raedler, T., Sommer, M., Sach, M., Weiller, C., and Koch, M. A. (2004), "White Matter Asymmetry in the Human Brain: A Diffusion Tensor MRI Study," *Cerebral Cortex*, 14, 945–951. [595]

Casella, G., and Berger, R. L. (2002), *Statistical Inference* (2nd ed.), Pacific Grove, CA: Duxbury. [592]

Chiu, T. Y. M., Leonard, T., and Tsui, K. W. (1996), "The Matrix-Logarithmic Covariance Model," *Journal of the American Statistical Association*, 91, 198–210. [589]

Dougherty, R. F., Ben-Shachar, M., Deutsch, G. K., Hernandez, A., Fox, G. R., and Wandell, B. A. (2007), "Temporal-Callosal Pathway Diffusivity Predicts Phonological Skills in Children," *Proceedings of the National Academy of Sciences of the USA*, 104, 8556–8561. [589,592]

Dougherty, R. F., Ben-Shachar, M., Deutsch, G. K., Potanina, P., and Wandell, B. A. (2005), "Occipital-Callosal Pathways in Children: Validation and Atlas Development," *Annals of the New York Academy of Sciences*, 1064, 98–112. [593]

Edelman, A., Arias, T. A., and Smith, S. T. (1998), "The Geometry of Algorithms With Orthogonality Constraints," *SIAM Journal on Matrix Analysis and Applications*, 20, 303–353. [597]

Efron, B. (2007), "Size, Power and False Discovery Rates," *The Annals of Statistics*, 35, 1351–1377. [595]

Fletcher, P. T., and Joshi, S. (2007), "Riemannian Geometry for the Statistical Analysis of Diffusion Tensor Data," *Signal Processing*, 87, 250–262. [589]

Flury, B. N. (1984), "Common Principal Components in  $k$  Groups," *Journal of the American Statistical Association*, 79, 892–898. [588]

- Genovese, C. R., Lazar, N. A., and Nichols, T. E. (2002), "Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate," *Neuroimage*, 15, 870–878. [589]
- Jones, D. K., Griffin, L. D., Alexander, D. C., Catani, M., Horsfield, M., Howard, R., and Williams, S. C. R. (2002), "Spatial Normalization and Averaging of Diffusion Tensor MRI Data Sets," *Neuroimage*, 17, 592–617. [588]
- Kuonen, D. (1999), "Saddlepoint Approximations for Distributions of Quadratic Forms in Normal Variables," *Biometrika*, 86, 929–935. [592]
- Lang, S. (1999), *Fundamentals of Differential Geometry*, New York: Springer-Verlag. [597]
- LeBihan, D., Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., and Chabriat, H. (2001), "Diffusion Tensor Imaging: Concepts and Applications," *Journal of Magnetic Resonance Imaging*, 13, 534–546. [588]
- Leonard, T., and Hsu, J. S. J. (1992), "Bayesian Inference for a Covariance Matrix," *The Annals of Statistics*, 20, 1669–1696. [589]
- Logan, B. R., and Rowe, D. B. (2004), "An Evaluation of Thresholding Techniques in fMRI Analysis," *Neuroimage*, 22, 95–108. [589]
- Mallows, C. L. (1961), "Latent Vectors of Random Symmetric Matrices," *Biometrika*, 48, 133–149. [589]
- Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979), *Multivariate Analysis*, San Diego, CA: Academic Press. [598]
- Mehta, M. L. (1991), *Random Matrices* (2nd ed.), San Diego, CA: Academic Press. [589]
- Moakher, M. (2002), "Means and Averaging in the Group of Rotations," *SIAM Journal on Matrix Analysis and Applications*, 24, 1–16. [597]
- Rauschecker, A. M., Deutsch, G. K., Ben-Shachar, M., Schwartzman, A., Perry, L. M., and Dougherty, R. F. (2009), "Reading Impairment in a Patient With Missing Arcuate Fasciculus," *Neuropsychologia*, 47, 180–194. [589]
- Schwartzman, A. (2006), "Random Ellipsoids and False Discovery Rates: Statistics for Diffusion Tensor Imaging Data," Ph.D. thesis, Stanford University. [589]
- (2008), "Empirical Null and False Discovery Rate Inference for Exponential Families," *The Annals of Applied Statistics*, 2, 1332–1359. [595]
- Schwartzman, A., Dougherty, R. F., Lee, J., Ghahremani, D., and Taylor, J. E. (2009), "Empirical Null and False Discovery Rate Analysis in Neuroimaging," *Neuroimage*, 44, 71–82. [589]
- Schwartzman, A., Dougherty, R. F., and Taylor, J. E. (2005), "Cross-Subject Comparison of Principal Diffusion Direction Maps," *Magnetic Resonance in Medicine*, 53, 1423–1431. [588]
- (2008), "False Discovery Rate Analysis of Brain Diffusion Direction Maps," *The Annals of Applied Statistics*, 2, 153–175. [588,596]
- Schwartzman, A., Mascarenhas, W., and Taylor, J. E. (2008), "Inference for Eigenvalues and Eigenvectors of Gaussian Symmetric Matrices," *The Annals of Statistics*, 36, 2886–2919. [588-590,595]
- Storey, J. D., Taylor, J. E., and Siegmund, D. (2004), "Strong Control, Conservative Point Estimation and Simultaneous Conservative Consistency of False Discovery Rates: A Unified Approach," *Journal of the Royal Statistical Society, Ser. B*, 66, 187–205. [596]
- Stuart, A., and Ord, J. K. (1994), *Kendall's Advanced Theory of Statistics* (6th ed.), London: Edward Arnold. [592]
- Whitcher, B., Wisco, J. J., Hadjikhani, N., and Tuch, D. S. (2007), "Statistical Group Comparison of Diffusion Tensors via Multivariate Hypothesis Testing," *Magnetic Resonance in Medicine*, 57, 1065–1074. [588,589]
- Worsley, K. J., Taylor, J. E., Tomaiuolo, F., and Lerch, J. (2004), "Unified Univariate and Multivariate Random Field Theory," *Neuroimage*, 23, S189–S195. [596]
- Wu, Y.-C., Field, A. S., Chung, M. K., Badie, B., and Alexander, A. L. (2004), "Quantitative Analysis of Diffusion Tensor Orientation: Theoretical Framework," *Magnetic Resonance in Medicine*, 52, 1146–1155. [588]
- Yao, Y. (1965), "An Approximate Degrees of Freedom Solution to the Multivariate Behrens Fisher Problem," *Biometrika*, 52, 139–147. [590]
- Zhu, H., Zhang, H., Ibrahim, J. G., and Peterson, B. S. (2007), "Statistical Analysis of Diffusion Tensors in Diffusion-Weighted Magnetic Resonance Imaging," *Journal of the American Statistical Association*, 102, 1085–1102. [589,596]