



## Sequential anisotropic multichannel Wiener filtering with Rician bias correction applied to 3D regularization of DWI data

M. Martin-Fernandez<sup>a,\*</sup>, E. Muñoz-Moreno<sup>a</sup>, L. Cammoun<sup>b</sup>, J.-P. Thiran<sup>b</sup>, C.-F. Westin<sup>c</sup>, C. Alberola-López<sup>a</sup>

<sup>a</sup>Laboratorio de Procesado de Imagen (LPI), Universidad de Valladolid, Valladolid, Spain

<sup>b</sup>Institut de Traitement des Signaux (ITS), Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

<sup>c</sup>Laboratory of Mathematics in Imaging (LMI), Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

### ARTICLE INFO

#### Article history:

Received 24 April 2007

Received in revised form 14 May 2008

Accepted 20 May 2008

Available online 7 June 2008

#### Keywords:

Minimum MSE criterion

Wiener filter

DWI data

Rice distribution

Bias correction

Parameter estimation

Sequential filter

Anisotropic filter

### ABSTRACT

It has been shown that the tensor calculation is very sensitive to the presence of noise in the acquired images, yielding to very low quality Diffusion Tensor Images (DTI) data. Recent investigations have shown that the noise present in the Diffusion Weighted Images (DWI) causes bias effects on the DTI data which cannot be corrected if the noise characteristic is not taken into account. One possible solution is to increase the minimum number of acquired measurements (which is 7) to several tens (or even several hundreds). This has the disadvantage of increasing the acquisition time by one (or two) orders of magnitude, making the process inconvenient for a clinical setting. We here proposed a turn-around procedure for which the number of acquisitions is maintained but, the DWI data are filtered prior to determining the DTI. We show a significant reduction on the DTI bias by means of a simple and fast procedure which is based on linear filtering; well-known drawbacks of such filters are circumvented by means of anisotropic neighborhoods and sequential application of the filter itself. Information of the first order probability density function of the raw data, namely, the Rice distribution, is also included. Results are shown both for synthetic and real datasets. Some error measurements are determined in the synthetic experiments, showing how the proposed scheme is able to reduce them. It is worth noting a 50% increase in the linear component for real DTI data, meaning that the bias in the DTI is considerably reduced. A novel fiber smoothness measure is defined to evaluate the resulting tractography for real DWI data. Our findings show that after filtering, fibers are considerably smoother on the average. Execution times are very low as compared to other reported approaches which allows for a real-time implementation.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

Regularization of Diffusion Tensor Images (DTI) is a problem of paramount importance as a prior step for finding connectivity in the brain. DTI are obtained from Diffusion Weighted Images (DWI), i.e., MR images acquired after applying gradients in different directions (Basser et al., 1994; Stejskal and Tanner, 1965). It has been shown that background noise present in DWI introduces artifacts in DTI (Basser and Pajevic, 2000). In particular, in Jones and Basser (2004) some of this artifacts were analyzed. This work reported that the background noise present in the DWI produces an underestimation (bias) of indices of diffusion anisotropy which is not possible to correct once the DTI data are determined. This work justifies the need of filtering the DWI before any attempt at

calculating the DTI data. In addition, in Basu et al. (2006) the authors presented some interesting results showing a decrease in the trace as well as a bias in the anisotropy.

Many attempts have been reported to tackle DTI regularization; (Tschumperle and Deriche, 2003) makes a good classification of the existing techniques. These approaches can be summarized as follows:

- DWI regularization techniques carry out DWI filtering prior to determining the DTI. In Parker et al. (2000), the authors choose a well-established method to perform non-linear anisotropic smoothing for each DW image independently. In particular, they used the method developed by Perona and Malik (1990). They do not incorporate dependencies among the DWI. In McGraw et al. (2004), the authors perform DWI smoothing by means of a total variation approach which can be casted as a variational technique in which a gradient descent method has to be solved. In addition, in Basu et al. (2006) the authors present a variational formulation to a maximum *a posteriori* estimation for the noise-free DWI, using a Rician maximum likelihood term and a prior smoothing term. This is performed independently for each DWI.

\* Corresponding author. Address: ETSIT, Universidad de Valladolid, Campus Miguel Delibes s/n, 47011 Valladolid, Spain. Tel.: +34 983 185551; fax: +34 983 423667.

E-mail addresses: [marcma@tel.uva.es](mailto:marcma@tel.uva.es) (M. Martin-Fernandez), [leila.cammoun@epfl.ch](mailto:leila.cammoun@epfl.ch) (L. Cammoun), [JP.Thiran@epfl.ch](mailto:JP.Thiran@epfl.ch) (J.-P. Thiran), [westin@bwh.harvard.edu](mailto:westin@bwh.harvard.edu) (C.-F. Westin), [caralb@tel.uva.es](mailto:caralb@tel.uva.es) (C. Alberola-López).

- For DTI estimation methods, a simultaneous regularization and estimation of DTI out of the DWI is performed. In Wang et al. (2004), the authors proposed a constrained variational framework for a direct estimation and regularization of DTI out of the complex DWI. The main drawback of this technique is that the complex DWI data are not always available but the natural envelope. A very interesting work (Fillard et al., 2007) has been recently published in which the DTI data are regularized using directly the DWI data: the DTI are simultaneously estimated and regularized. In this work several models for the DWI are compared. The Rice model for the DWI gives rise to the better performance. The problem is posed as a maximum *a posteriori* estimation under a variational framework. In addition, the authors make use of the Riemannian approach and Log-Euclidean metrics proposed in Pennec et al. (2006). In this approach, the space spanned by positive definite tensors is replaced by a regular and complete manifold for which null eigenvalues are at the infinity and negative values are not allowed in the corresponding tensor space. Common image processing techniques as interpolation and Gaussian smoothing can be thus performed unrestrictedly throughout this new space. The tensor is constructed by projecting points from the manifold back onto the tensor space. To our knowledge and in our opinion, this is the best approach to solve the noise problem in DTI.
- On the other hand, in DTI regularization approaches, the filtering is directly performed on the noisy DTI. Constraints of positive semi-definiteness are usually imposed. Some techniques work directly on the tensor elements (Chefd'hotel et al., 2002; Martin-Fernandez et al., 2004; Pennec et al., 2006) while others do it on the spectral tensor decomposition (Poupon et al., 2000; Coulon et al., 2004; Chen and Hsu, 2005). In Chefd'hotel et al. (2002), the authors propose a non-linear partial differential approach constrained to lie on a manifold on which the tensors are positive semi-definite. Martin-Fernandez et al. (2004) poses the problem as a Bayesian estimation problem making use of a Markov Random Field and multivariate Gaussian distributions. Concerning the second approach, in Poupon et al. (2000) an energy function based on a Markovian model was used to regularized the noisy dominant eigenvector. In Coulon et al. (2004), an iterative restoration scheme for principal diffusion direction is proposed. Finally, in Chen and Hsu (2005) a vector extension of the anisotropic filtering of Perona and Malik (1990) is developed to smooth the tensor eigenvectors.

Most of these works are numerically intensive and do not make use of the underlying statistical distribution. The work by Basu et al. (2006) is very interesting as it made use of the Rice model, but the DWI were processed independently which, as pointed out in Fillard et al. (2007) is not a good idea: “only the combination of all the DWI reveals the complex neural structure of the white matter”. So a method should model the DWI jointly and exploit the existing correlation among the DW components. The method by Fillard et al. (2007) is, in our opinion, very interesting, as it has all the good ingredients. However, it does not exploit the correlation among the signals in the DWIs, as the joint probability function is built by multiplying the marginals which means the DWIs are considered as independent signals. In addition, it is computationally expensive and could not be considered as real time: the execution time reported was 12 min. In this paper, we present a fast and simple method with results similar to the work by Fillard et al. to filter the DWI prior to the calculation of the DTI. It is based on a modified Wiener filter. It exploits the whole correlations among the DWIs. Drawbacks of this approach<sup>1</sup> which are

well known and have been comparatively highlighted elsewhere (for instance, in Chen and Hsu, 2005) will be circumvented by two main ideas. First, information about the probabilistic nature of DWI data (i.e. the Rice distribution Gudbjartsson and Patz, 1995) will be accounted for. Second, anisotropy in the filter will be included by means of estimating the parameters of the filter within estimation windows whose shapes are a function of the local characteristics of the data. Furthermore, the filter is applied sequentially to refine such estimates. Additional assumptions let us formulate the problem in closed-form manageable expressions. We come up with a simple and fast procedure which yields very good results both on synthetic images and on real DWI. In addition, the filter allows an important reduction of the DTI bias reported in the literature. A version of this filter dealing with scalar Magnetic Resonance signals (which does not include information of the Rice distribution of the data) has been recently published (Martin-Fernandez et al., 2007) and has drawn results comparable to other schemes based on the diffusion equation, at a lower computational cost. Here we extend this work in the direction indicated above.

## 2. Methods

### 2.1. Isotropic solution

Let  $\mathbf{Y}$  denote the observed dataset defined for a 3D grid with size  $S = P \times Q \times R$ . The DW dataset contains the six scalar volumes for the six diffusion gradient directions as well as the baseline data. The total number of variables in the observed dataset is  $7S$ . This dataset will be considered as a column vector created by stacking column vectors of seven components (the observations at every spatial position, each from every gradient used); the stack is created by visiting each voxel with some predefined order, but the order itself is irrelevant to the purposes of this paper.

The model  $\mathbf{Y} = \mathbf{X} + \mathbf{N}$  is assumed, where  $\mathbf{X}$  and  $\mathbf{N}$  are the signal and the noise components, respectively. The filter performs the following operation<sup>2</sup>:

$$\mathbf{Z} = \mathbf{C}_X[\mathbf{C}_X + \mathbf{C}_N]^{-1}[\mathbf{Y} - \boldsymbol{\eta}_X] + \boldsymbol{\eta}_X, \quad (1)$$

where  $\boldsymbol{\eta}_X$  and  $\mathbf{C}_X$  are the signal means and covariance matrix, respectively, and  $\mathbf{C}_N$  is the covariance of the noise.

$S$  is usually very large so the operation above is very involved; in addition, correlations tend to diminish for voxels sufficiently distant. An extreme case is to assume that voxels are mutually uncorrelated, both for the signal and for the noise. If this is the case, the covariance matrices  $\mathbf{C}_X$  and  $\mathbf{C}_N$  would be block diagonal, i.e., only  $S$  blocks with size  $7 \times 7$  located in the main diagonal would be non-zero. Specifically, denoting with  $\mathbf{Y}(p)$  the vector of the seven measurements for voxel  $p$ , with  $1 \leq p \leq S$ , the covariance matrix of this vector would be  $\mathbf{C}_X(p, p) + \mathbf{C}_N(p, p)$ , and this matrix would be the  $p$ th block in the main diagonal of matrix  $\mathbf{C}_X + \mathbf{C}_N$ ; it turns out that the inverse of this matrix is also block diagonal, where each block  $p$  in the main diagonal equals the inverse of matrix  $\mathbf{C}_X(p, p) + \mathbf{C}_N(p, p)$ . This result makes Eq. (1) be decoupled voxelwise as

$$\mathbf{Z}(p) = \mathbf{U}\{\mathbf{C}_X(p, p)[\mathbf{C}_X(p, p) + \mathbf{C}_N(p, p)]^{-1}[\mathbf{Y}(p) - \boldsymbol{\eta}_X(p)] + \boldsymbol{\eta}_X(p)\}. \quad (2)$$

The vector function  $\mathbf{U}\{\cdot\}$  is defined for a vector variable. It sets each negative element of the input vector to zero. This is needed in order to guarantee that the solution lies in the positive hyper-quadrant of the seven-dimensional space. This correction keeps the positive property of envelopes unaltered and it allows the Stejskal–Tanner

<sup>1</sup> Such as the Gaussianity assumption for the data and the fact that the filter leaves the data unaltered at boundaries. In addition, the Wiener filter is largely affected by errors in the estimation when true parameters are unknown.

<sup>2</sup> For the derivation and the assumptions of this expression, see Appendix A. Further details can also be found in Martin-Fernandez et al. (2007).

equation (Stejskal and Tanner, 1965) to be solved properly, as it was the case with the original data. However, this does not insure that the estimated tensors have always positive eigenvalues. In practice, this is not a problem. We will further discuss on that in Section 4. We can say that the local filter is linear, memoryless and space-variant. It is worth noting that the assumed model both for the signal and for the noise, although simplified, is still non-stationary as both mean vectors and covariance matrices are functions of the voxel position  $p$ . To assume that the signal is stationary is far too restrictive, but this assumption is reasonable for the noise. Accepting this, we can drop index  $p$  from the noise covariance matrix, yielding<sup>3</sup>

$$\mathbf{Z}(p) = \mathbf{U}\{\mathbf{C}_x(p)[\mathbf{C}_x(p) + \mathbf{C}_N]^{-1}[\mathbf{Y}(p) - \boldsymbol{\eta}_x(p)] + \boldsymbol{\eta}_x(p)\}. \quad (3)$$

In addition, as the DW images are acquired independently, the noise samples present in two such images can be assumed uncorrelated. For this reason, we will assume the noise covariance matrix  $\mathbf{C}_N$  diagonal. We have to consider only seven noise variances  $\sigma_N^2(k)$  with  $0 \leq k \leq 6$ . The number of unknowns to estimate has been considerably reduced; this number is equal to  $35S + 7$ :  $7S$  signal means,  $28S$  signal covariance parameters<sup>4</sup>, and seven noise variances. If these parameters were known, Eq. (3) would give the local Wiener solution; however, and generally speaking, this is not usually the case so we have to find out how to estimate these parameters in advance from the observation  $\mathbf{Y}$ .

We can rely on some assumption about the signal  $\mathbf{X}$ . Though we said that this signal is clearly non-stationary, it is reasonable to assume that the local variation of both the means and the autocovariance matrices within a given neighborhood is less than the global variation, i.e., we will assume that the signal has only local interactions (Besag, 1974). Our assumption is that the signal can be considered locally ergodic so as to use spatial averages to estimate the unknown parameters.

Let  $\boldsymbol{\vartheta}(p)$  denote a set of  $L$  neighbors corresponding to the site (voxel)  $p$ . The set  $\boldsymbol{\vartheta}(p)$  also includes the voxel  $p$ . We will bear in mind neighborhoods of  $L = |\boldsymbol{\vartheta}(p)| = 27$  voxels.

We can use those neighborhoods  $\boldsymbol{\vartheta}(p)$  to estimate the  $35S + 7$  unknowns. As the neighborhoods are isotropic this filter will be referred to as *isotropic Wiener*. This is the first approach we will present. Later on we will introduce an anisotropic counterpart.

Assuming local ergodicity in the set  $\boldsymbol{\vartheta}(p)$  for site  $p$ , we get estimates

$$\boldsymbol{\eta}_x(p) = \boldsymbol{\eta}_y(p) \approx \bar{\mathbf{Y}}(p) = \frac{1}{|\boldsymbol{\vartheta}(p)|} \sum_{p' \in \boldsymbol{\vartheta}(p)} \mathbf{Y}(p'), \quad (4)$$

$$\begin{aligned} \mathbf{C}_x(p) &\approx \mathbf{C}_y(p) \approx \bar{\mathbf{C}}_y(p) \\ &= \frac{1}{|\boldsymbol{\vartheta}(p)| - 1} \sum_{p' \in \boldsymbol{\vartheta}(p)} (\mathbf{Y}(p') - \bar{\mathbf{Y}}(p'))(\mathbf{Y}(p') - \bar{\mathbf{Y}}(p'))^T. \end{aligned} \quad (5)$$

The estimation of the noise variances  $\sigma_N^2(k)$ , for  $0 \leq k \leq 6$ , is more involved and will have a great influence in the filter output as we will see later. The problem is due to the fact that the autocovariance matrices add up as  $\mathbf{C}_y(p) = \mathbf{C}_x(p) + \mathbf{C}_N$ . One possible solution is to look for a region in the image grid for which the signal components are *a priori* known to be zero. In this case  $\mathbf{C}_x(p)$  will be zero and we will have that  $\sigma_N^2(k)$  will be equal to the element  $(k, k)$  of the autocovariance matrix  $\mathbf{C}_y(p)$  in that region. Thus, the sample variances in that region would give us a reasonable estimation of the noise variances. Nevertheless, that piece of knowledge may not be available or it would require the user's supervision. The proposed solution is to introduce a free parameter  $\lambda$  ranging

in the interval  $(0, 1)$  that will be hereafter referred to as *regularization parameter* which would allow us to fine tune the amount of regularization required by means of a selection of the noise variances between two given bounds as

$$\bar{\sigma}_N^2(k) = (1 - \lambda)\bar{\sigma}_{\min}^2(k) + \lambda\bar{\sigma}_{\text{ave}}^2(k), \quad (6)$$

where  $\bar{\sigma}_{\min}^2(k)$  is the  $(k, k)$  element of the the autocovariance matrix  $\bar{\mathbf{C}}_y(p')$  with

$$p' = \arg \min_{1 \leq p \leq S} \text{TR}[\bar{\mathbf{C}}_y(p)], \quad (7)$$

where  $\text{TR}[\cdot]$  stands for the matrix trace and  $\bar{\sigma}_{\text{ave}}^2(k)$  is the  $(k, k)$  element of

$$\frac{1}{S} \sum_{p=1}^S \bar{\mathbf{C}}_y(p). \quad (8)$$

We can say the higher the  $\lambda$  value, the higher the noise variances  $\bar{\sigma}_N^2(k)$  and, consequently, the higher the regularization and vice versa. This parameter may be either manually set or learned by training for a specific target application. Consequently, in Eq. (3) we can replace the parameters by their corresponding estimates giving rise to

$$\mathbf{Z}(p) \approx \mathbf{U}\{\bar{\mathbf{C}}_y(p)[\bar{\mathbf{C}}_y(p) + \bar{\mathbf{C}}_N]^{-1}[\mathbf{Y}(p) - \bar{\mathbf{Y}}(p)] + \bar{\mathbf{Y}}(p)\}, \quad (9)$$

where  $\bar{\mathbf{C}}_N$  is a diagonal matrix whose diagonal elements are the noise variance estimates  $\bar{\sigma}_N^2(k)$ .

Eqs. (4)–(6) are coarse approximations, thus a direct implementation of Eq. (9) will lead to poor results, therefore, we propose here a *sequential* scheme (Hillery and Chin, 1991), as follows:

$$\mathbf{Y}^{(n+1)}(p) = \bar{\mathbf{C}}_y^{(n)}(p)[\bar{\mathbf{C}}_y^{(n)}(p) + \bar{\mathbf{C}}_N^{(n)}]^{-1}[\mathbf{Y}^{(n)}(p) - \bar{\mathbf{Y}}^{(n)}(p)] + \bar{\mathbf{Y}}^{(n)}(p) \quad (10)$$

with the initial condition

$$\mathbf{Y}^{(0)}(p) = \mathbf{Y}(p). \quad (11)$$

The estimates are recalculated at iteration  $(n)$  using Eqs. (4)–(6) with the data  $\mathbf{Y}^{(n)}(p)$ . A number of iterations  $N \in (5, 10)$  has been considered in the experiments as sufficient. The final solution is then given by

$$\mathbf{Z}(p) \approx \mathbf{U}\{\mathbf{Y}^{(N)}(p)\}. \quad (12)$$

This will also provide local coherence among neighboring voxels, a fact that palliates the (somewhat restrictive) assumption of uncorrelation within the signal components, i.e., the sequential approach is close to a non-sequential version in which the voxels are assumed to be correlated.

## 2.2. Anisotropic solution

The Wiener filter introduced so far can be consider as isotropic. Hereafter we will describe how we can modify the estimation proposed in Eqs. (4)–(6) to be considered as anisotropic. This approach will be referred to as *anisotropic Wiener*. In Fig. 1, we can see six different non-homogeneous neighborhood subsystems  $\boldsymbol{\vartheta}_m(p)$ , with  $1 \leq m \leq 6$ , for which a specific orientation is assumed. We can repeat Eqs. (4) and (5) for each voxel and each orientation  $m$  using the subsystem  $\boldsymbol{\vartheta}_m(p)$  to obtain oriented means

$$\bar{\mathbf{Y}}^m(p) = \frac{1}{|\boldsymbol{\vartheta}_m(p)|} \sum_{p' \in \boldsymbol{\vartheta}_m(p)} \mathbf{Y}(p') \quad (13)$$

and oriented autocovariance matrices

$$\bar{\mathbf{C}}_y^m(p) = \frac{1}{|\boldsymbol{\vartheta}_m(p)| - 1} \sum_{p' \in \boldsymbol{\vartheta}_m(p)} (\mathbf{Y}(p') - \bar{\mathbf{Y}}^m(p'))(\mathbf{Y}(p') - \bar{\mathbf{Y}}^m(p'))^T. \quad (14)$$

<sup>3</sup> For the sake of simplicity, we have dropped the second  $p$  index for the signal covariance matrix as  $\mathbf{C}_x(p, p) = \mathbf{C}_x(p)$ .

<sup>4</sup> A  $7 \times 7$  symmetric autocovariance matrix has 28 degrees of freedom, thus we have 28 signal covariance parameters per voxel and  $S$  voxels.

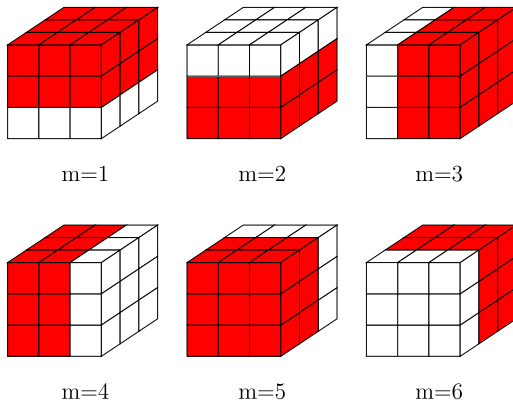


Fig. 1. Oriented neighborhood subsystems with  $L = 27$  encoding six different orientations (red sites).

Then the true orientation  $q(p)$  can be estimated as

$$q(p) \approx \arg \min_{1 \leq m \leq 6} \text{TR}[\overline{\mathbf{C}}_Y^m(p)]. \quad (15)$$

The final estimation for  $\eta_X(p) = \eta_Y(p)$  and  $\mathbf{C}_X(p) \approx \mathbf{C}_Y(p)$  can be determined by using Eqs. (13) and (14) with  $m = q(p)$  for each position  $p$ , respectively. Thus, we write

$$\eta_X(p) = \eta_Y(p) \approx \overline{Y}^{q(p)}(p) \quad \text{and} \quad \mathbf{C}_X(p) \approx \mathbf{C}_Y(p) \approx \overline{\mathbf{C}}_Y^{q(p)}(p). \quad (16)$$

Then, Eq. (6) can be used to estimate the noise variances  $\overline{\sigma}_N^2(k)$  as explained before. We also propose to implement a sequential version of the anisotropic filter using the new anisotropic parameter estimation method by using Eqs. (10)–(12) as explained in Section 2.1.

### 2.3. Non-Gaussian corrections: inclusion of the Rice distribution

As it is well known, the Wiener filter, as explained along previous sections, only makes use of first and second order statistics, since this is the only information used by the linear minimum MSE estimator. As shown in Appendix B, the noise-free natural envelope  $s$  of the signal (Eq. (B.6)) is very far from the expected value of the noisy natural envelope (Eq. (B.9)) for low signal to noise ratios. This will introduce a bias in the filtering procedure, since the solution is always centered about the signal mean as pointed out by Eq. (3). We need to estimate the Rice parameter  $s$  in order to compensate for the bias. However the noise parameter  $\sigma$  is also unknown. Maximum likelihood of both parameters is possible, but no closed-form solution for it exists (Sijbers and den Dekker, 2004). An iterative algorithm is needed, but it would be computationally very expensive. It will not give rise to a fast implementation. We have resorted the estimation problem to a simple and fast solution which follows.<sup>5</sup>

The bias effect has to be considered mainly for low SNR. For high SNR, as the Rice distribution approaches a Gaussian distribution, the Wiener solution is the general optimum solution with no bias. In Fig. 2, we can see the SNR as a function of the coherent to non-coherent ratio, i.e., the  $\gamma$  parameter. We referred to this function as  $B(\gamma)$  function. See Appendix B for its derivation.

The SNR can be estimated by the method of moments as

$$\text{SNR} \approx \overline{\text{SNR}} = \frac{\overline{Y}}{\sqrt{Y^2 - \overline{Y}^2}}. \quad (17)$$

<sup>5</sup> The basic material – including definitions of the parameters involved – concerning the Rice distribution has been written in Appendix B.

From Fig. 2, it is clear that the  $B(\cdot)$  function is a monotonic function of  $\gamma$ , thus the inverse function  $B^{-1}(\cdot)$  is well defined. We can write

$$\gamma = B^{-1}(\text{SNR}). \quad (18)$$

This inverse function is shown in Fig. 3. An estimate of the  $\gamma$  parameter can be determined by replacing the SNR value by its estimate given by Eq. (17). We can write then

$$\gamma \approx \bar{\gamma} = \begin{cases} B^{-1}(\overline{\text{SNR}}) & \overline{\text{SNR}} > \sqrt{\frac{\pi}{4-\pi}} \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

We can write the mean-squared value using  $\gamma$  as

$$E\{Y^2\} = s^2 \frac{2 + \gamma^2}{\gamma^2}, \quad \text{therefore } s = \sqrt{\frac{E\{Y^2\} \gamma^2}{2 + \gamma^2}}. \quad (20)$$

Finally, by replacing the parameters by their estimates, we can write

$$\bar{s} = \sqrt{\frac{\overline{Y^2} \bar{\gamma}^2}{2 + \bar{\gamma}^2}}. \quad (21)$$

The bias correction is performed by means of

$$Y'(p) = \begin{cases} Y(p) - \overline{Y}(p) + \bar{s} & Y(p) - \overline{Y}(p) + \bar{s} > 0, \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

for each voxel  $p$ . This is done independently for each of the seven DW images. Eq. (22) constrains the output to be positive, as an envelope cannot be negative. Negative values for the DWI are not possible as a logarithm is used to calculate the DTI (see for instance Eq. (6) in Westin et al., 2002). After this bias correction step, the procedures defined in Sections 2.1 and 2.2 will be applied to the modified dataset. These methods are also constrained to be within the positive hyper-quadrant in the seven-dimensional space. In conclusion, no negative values are permitted so that the DTI data can be computed without difficulties as it was the case for the original data. However, in principle this does not guarantee the positive-semi-definiteness property in the estimated tensors (see Section 4). However, the smoothing introduced with the use of the neighboring information improves the quality of the estimation as it will be shown in the experiments.

### 2.4. Synthetic data

We have created three  $50 \times 50 \times 50$  DT synthetic images with different shapes (namely, Earth, logarithm and cross data sets, see Appendix C for a description about how these tensor data sets have been created) to evaluate the performance of the framework proposed. The associated eigenvalues for the Earth and logarithm were  $(\lambda_1, \lambda_2, \lambda_3) = (7, 2, 1)$ . The final tensor field was normalized by  $10^{-4}$ . Using a value of  $b = 1000$  and the six gradient directions  $(1, 1, 0)$ ,  $(0, 1, 1)$ ,  $(1, 0, 1)$ ,  $(0, 1, -1)$ ,  $(-1, 1, 0)$  and  $(-1, 0, 1)$  the tensor field was converted to DWI data, using as baseline the trace of the tensors. Rice noise with  $\text{SNR} = 10$  was added to form the noisy dataset. Finally, the noisy tensor field was determined. The stochastic fiber tracking method (the random walk algorithm) proposed in Hagmann et al. (2003) was used.

The cross synthetic data have a crossing region where two orthogonal fibers cross. This region have planar anisotropy. This example was selected to show that the filter is able to regularize regions with uncertainty. Algorithms filtering only the principal direction diffusion will fail in this case. The stochastic fiber tracking algorithm should be able to follow both paths in a smooth manner. The Earth example was chosen to see whether a tensor field with a high variation in orientation is properly filtered. In

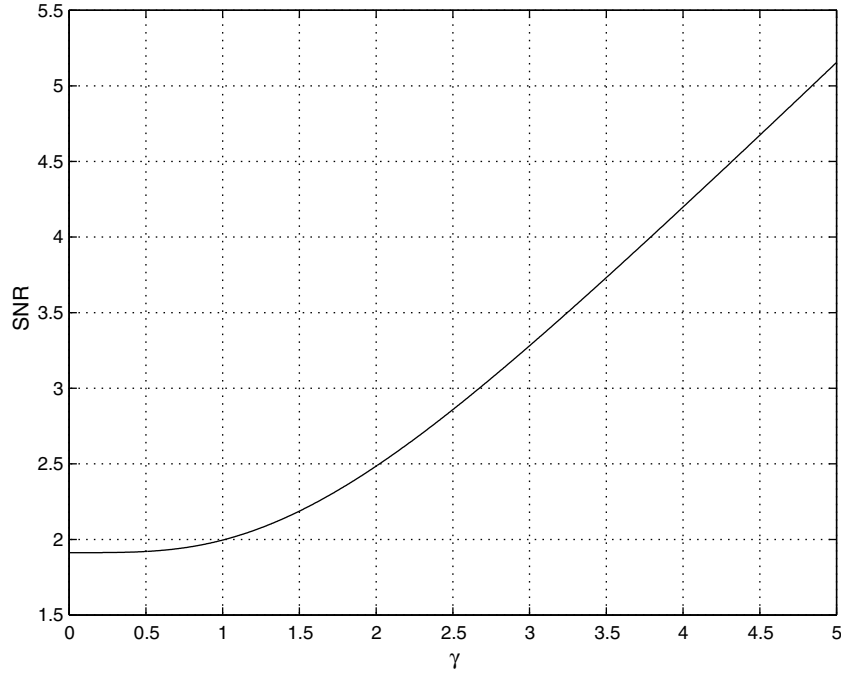


Fig. 2. SNR as a function of the  $\gamma$  parameter.

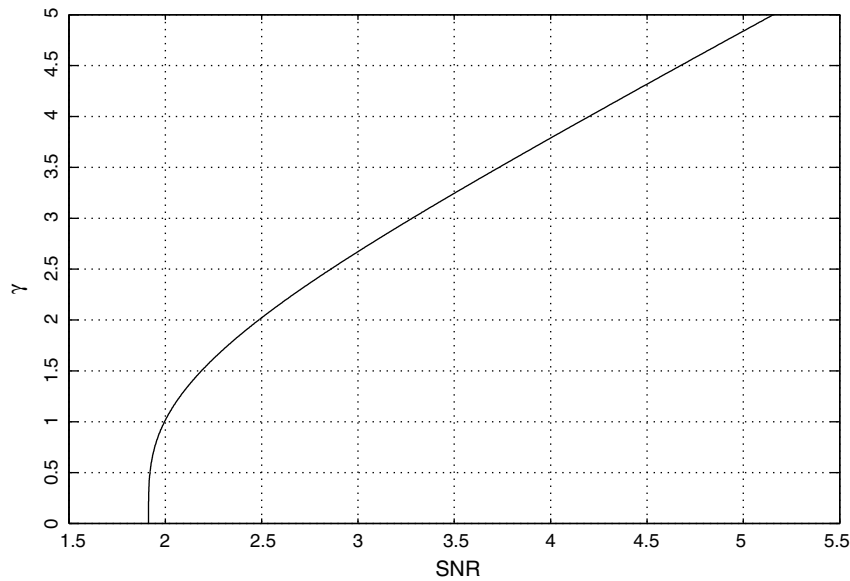


Fig. 3.  $\gamma$  parameter as a function of the SNR.

addition, in this example the boundaries among the regions are not straight and allows to test whether the anisotropic part of the filter works properly, even though the anisotropic parameter estimation uses the simple boundaries shown in Fig. 1. Finally, the logarithm example was selected to measure the averaged change in the anisotropy (bias correction). It is a tensor field having only prolate tensors allowing some orientation variation.

We define different measures to evaluate the performance of the filters. The DWI filtering error  $\epsilon$  can be defined as

$$\epsilon = \mathbf{Z} - \mathbf{X}. \quad (23)$$

The squared bias of the error is defined by

$$\text{Bsqr} = \frac{1}{49S^2} \left( \sum_{k=1}^7 \sum_{p=1}^S \epsilon_k(p) \right)^2, \quad (24)$$

where  $\epsilon_k(p)$  is the error at  $p$ th voxel (out of  $S$ ) in the  $k$ th acquisition image (out of 7). The mean-squared error and the variance are defined by

$$\text{MSE} = \frac{1}{7S} \sum_{k=1}^7 \sum_{p=1}^S \epsilon_k(p)^2 \quad \text{and} \quad \text{Var} = \text{MSE} - \text{Bsqr}. \quad (25)$$

In order to take into account the data variability,  $M$  noisy samples under the same scenario have been generated. We can then calculate statistics at each voxel position for the DTI data. Specifically, we define the total variance  $\text{TV}(p)$  of a diffusion tensor  $\mathbf{D}(p)$  at voxel position  $p$ , with  $1 \leq p \leq S$ , as

$$\text{TV}(p) = \text{TR}[\mathbf{C}_D(p)], \quad (26)$$

where for each voxel  $p$ ,  $\mathbf{C}_D(p)$  is the covariance matrix of the components of tensor  $\mathbf{D}(p)$  as

$$\mathbf{C}_D(p) = E\{\mathbf{D}_V(p)\mathbf{D}_V(p)^T\} - E\{\mathbf{D}_V(p)\}E\{\mathbf{D}_V(p)\}^T \quad (27)$$

and  $\mathbf{D}_V(p)$  is a column vector consisting of the six different elements of the symmetric tensor  $\mathbf{D}(p)$  arranged in any predefined order.

In Eq. (27), the expectations are replaced by sample means using the  $M$  available samples per voxel. In this case the sample total variance  $\overline{\text{TV}}(p)$  will be given by

$$\overline{\text{TV}}(p) = \text{TR}\{\overline{\mathbf{C}}_D(p)\}, \quad (28)$$

where for each voxel  $p$ ,  $\overline{\mathbf{C}}_D(p)$  is the sample covariance matrix defined as

$$\overline{\mathbf{C}}_D(p) = \frac{1}{M} \sum_{m=1}^M \mathbf{D}_V^{(m)}(p)\mathbf{D}_V^{(m)}(p)^T - \frac{1}{M^2} \left( \sum_{m=1}^M \mathbf{D}_V^{(m)}(p) \right) \left( \sum_{m=1}^M \mathbf{D}_V^{(m)}(p)^T \right) \quad (29)$$

and  $\mathbf{D}_V^{(m)}(p)$  is the  $m$ th sample of a column vector arrangement for sample tensor  $\mathbf{D}^{(m)}(p)$ .

In addition to previous measures, Bassler (1997) and Jones (2003) propose an interesting measure to quantify DTI uncertainty due to noise. They propose a new glyph representation referred to as “cone of uncertainty”, i.e., a cone whose cone angle is equal to the uncertainty (i.e., a given confidence interval) in the estimate of the orientation of the principal eigenvector and with a long axis that coincides with the mean eigenvector. We follow this idea to estimate the uncertainty.

In order to measure the confidence interval (uncertainty), we have simulated  $M$  noisy data sets. Let assume  $\mathbf{v}_1^{(m)}(p)$  is the eigenvector associated to the largest eigenvalue for the  $m$ th sample tensor  $\mathbf{D}^{(m)}(p)$  at voxel position  $p$ . Let  $\mathbf{v}_1(p)$  denote this same vector but of the ground truth tensor  $\mathbf{T}(p)$  at voxel position  $p$ . The angle between these two vectors is given by

$$\alpha^{(m)} = \arccos(|\mathbf{v}_1^{(m)}(p)^T \mathbf{v}_1(p)|). \quad (30)$$

In this expression the modulus allows to account for the problem of antipodal symmetry, i.e., eigenvectors are only defined up to its orientation along a particular axis. The closer to zero the  $\alpha$  value the lower the uncertainty. Since in this case, Eq. (30) is always greater than zero, the distribution of  $\alpha$  is necessarily one-sided. Hence, the value of  $\alpha^{(m)}$  at the 0.95Mth position in the list of  $M$  samples (sorted in ascending order) constitutes the upper bound of the 95% confidence interval (CI) of the angles, i.e., the uncertainty.

In addition to this measurements, and to measure the anisotropy, we use the Fractional Anisotropy (FA) value as well as the linear, planar and spherical components as defined in Westin et al. (2002). With respect to the orientation, it is color coded.<sup>6</sup> We use elliptic glyphs to represent some of the results.<sup>7</sup>

In addition to that, and to show the anisotropic performance of the filter, we have added an additional experiment comparing the isotropic solution with the anisotropic one. We have simulated a new synthetic data set as explained above with two clear discontinuities having only prolate tensors with orthogonal orientations.

### 2.5. Real imaging data

A DWI volume dataset using single-shot spin Echo Planar Imaging (EPI) sequence with diffusion-tensor encoding using six directions  $(1, 1, 0)$ ,  $(0, 1, 1)$ ,  $(1, 0, 1)$ ,  $(0, 1, -1)$ ,  $(1, -1, 0)$  and  $(-1, 0, 1)$ , size  $128 \times 128 \times 54$  and voxel size  $1.86 \times 1.86 \times 2.0$  mm was acquired. The non-biased anisotropic Wiener filter was run sequen-

<sup>6</sup> Red means left–right, green anterior–posterior and blue inferior–superior. The color intensity is scaled by the FA value to avoid coloring isotropic regions.

<sup>7</sup> Glyphs are cropped in order to avoid overlapping whenever the tensor trace is too large. This often happens in isotropic regions such as the ventricles.

tially with  $N = 5$  iterations and  $\lambda = 0.5$ . DTI datasets were determined both for the observed data and for the resulting data. Anisotropy and orientation are determined out of the tensor field. Additionally, we show a glyph representation for three manually selected regions of interest. In order to quantify the bias correction in the DTI, we have segmented for each region of interest the sites corresponding to left–right (red), anterior–posterior (green) and inferior–superior (blue) by thresholding the corresponding channel in the RGB color giving rise to nine subregions: three new regions within each manually selected region of interest. Then, we have determined the averaged linear component, as defined in Westin et al. (2002), in each subregion before and after filtering.

Finally, the stochastic method proposed in Hagmann et al. (2003) has been applied to determine the fiber tracts. A quantitative measure  $\mu$  of the smoothness along fibers has been calculated as explained in Appendix D. This parameter was defined so that the lower the  $\mu$  value, the smoother the fiber. An averaged smoothness  $\bar{\mu}$  is defined for a set of fibers.

## 3. Results

### 3.1. Synthetic data

One noise-free synthetic tensor data set was generated for the Earth, logarithm and cross cases. Tractographies for the three noise-free synthetic data sets are shown in Fig. 4. In particular, Fig. 4a shows a tractography for the noise-free Earth data set, Fig. 4b for the logarithm and Fig. 4c for the cross.

We have evaluated both the biased version of the filter (the one filtering on data  $Y$ ) and the non-biased version (the one that filters the data  $Y'$  – see Eq. (22)) for each noisy DWI synthetic data set. Additionally, the parameters were estimated both on the whole neighborhood (with  $L = 27$  neighbors) and on the oriented neighborhoods (as indicated in Fig. 1 and Eq. (15)). Squared bias (Bsq), variance (Var) and MSE for the filtered DWI data sets (see Eqs. (24) and (25)) were calculated. Achieved results are shown in Tables 1–4. These tables show results for the three synthetic examples (columns): cross, logarithm and Earth. Values in the tables were scaled by a factor of  $10^8$  for the sake of clarity. Minimum attained values have been highlighted for each measurement and synthetic case.

Table 1 represents the values for the observed noisy data. Table 2 shows values of the estimation of the parameter  $s$  only prior to the filtering stage as explained in Section 2.3. Two type of estimation schemes were used (isotropic and anisotropic) as indicated by the last column (as described in Sections 2.1 and 2.2, respectively).

Table 3 shows results after applying the filter for  $\lambda = 0.5$  and  $N = 5$  iterations. Isotropic and anisotropic versions of the filter as well as with and without bias correction are presented and ordered as indicated by the last column. Table 4 gives similar results but for  $N = 10$  iterations.

$M = 100$  noisy synthetic DWI data sets were computed for the logarithm and the Earth cases. Each noisy data set was then filtered. The anisotropic version of the filter with bias correction,  $N = 5$  iterations and  $\lambda = 0.5$  was selected for this experiment. The corresponding DTI data sets for the  $M$  executions were computed. The estimated total tensor variance per voxel was then determined using Eqs. (28) and (29) for the noisy and for the filtered DTI datasets. Fig. 5 shows the achieved distribution of the total variance for the different voxel positions. In particular, Fig. 5a shows the result for the noisy data sets and Fig. 5b for the filtered ones in the Earth case.

A CI of the uncertainty angle (the angle of the “cone of uncertainty”) per voxel was determined as explained in Section 2.4. Fig. 6 shows the achieved distribution of the CIs for the different

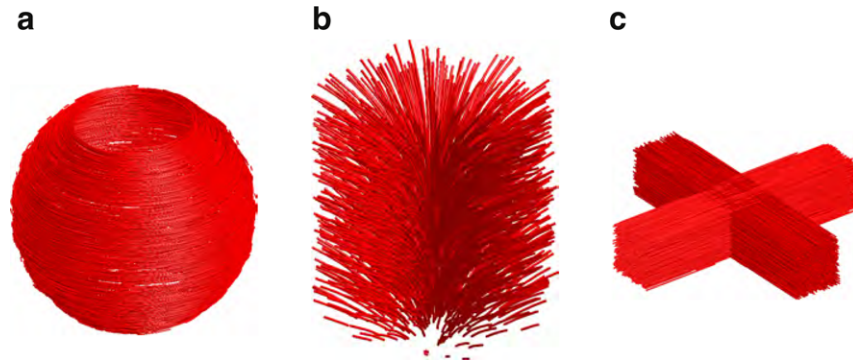


Fig. 4. Tractography for the synthetic images: (a) Earth, (b) logarithm and (c) cross.

Table 1

Original data corrupted by noise: bias, variance and MSE

Measure	Cross	Logarithm	Earth
Bsq	0.0309	0.1376	0.0824
Var	0.8357	3.8455	1.5228
MSE	0.8666	3.9831	1.6052

Table 2

Results for the estimation of the  $s$  parameter with no filtering

Measure	Cross	Logarithm	Earth	Estimator type
Bsq	0.0088	<b>0.0028</b>	0.0885	Isotropic
Var	1.1576	4.0046	2.3811	
MSE	1.1664	4.0074	2.4695	
Bsq	<b>0.0027</b>	0.0146	<b>0.0003</b>	Anisotropic
Var	<b>0.8537</b>	<b>3.8704</b>	<b>1.7208</b>	
MSE	<b>0.8564</b>	<b>3.8850</b>	<b>1.7211</b>	

Minimum attained values highlighted for each measurement and synthetic case.

Table 3

Results after filtering for  $\lambda = 0.5$  and  $N = 5$  iterations

Measure	Cross	Logarithm	Earth	Filter type
Bsq	0.0311	0.1387	0.0839	Isotropic
Var	<b>0.0302</b>	0.2288	0.1024	
MSE	0.0613	0.3675	0.1863	
Bsq	0.0093	<b>0.0027</b>	0.1038	No-bias isotropic
Var	0.2778	0.2734	0.5659	
MSE	0.2871	0.2762	0.6697	
Bsq	0.0286	0.1321	0.0755	Anisotropic
Var	0.0397	0.2055	<b>0.0791</b>	
MSE	0.0683	0.3376	<b>0.1546</b>	
Bsq	<b>0.0023</b>	0.0138	<b>0.0001</b>	No-bias anisotropic
Var	0.0525	<b>0.2055</b>	0.1589	
MSE	<b>0.0548</b>	<b>0.2193</b>	0.1590	

Minimum attained values highlighted for each measurement and synthetic case.

voxel position. In particular, Fig. 6a shows the result for the noisy data sets and Fig. 6b for the filtered ones in the logarithm case.

For the total variance and the CI of the uncertainty angle, and for each case, mean and standard deviation were computed for the noisy and the filtered values. In addition, an unbalanced one-sided  $t$ -test with unknown unequal variances at a significance level of  $\alpha = 0.05$  was also performed to check whether or not the mean of the total variances and the mean CI of the uncertainty angle measures for the filtered valued is smaller than the noisy counterpart. Results are shown in Table 5. Mean and standard deviation for the total variance was scaled by a factor of  $10^8$  for the sake of clarity.

Table 4

Results after filtering for  $\lambda = 0.5$  and  $N = 10$  iterations

Measure	Cross	Logarithm	Earth	Filter type
Bsq	0.0314	0.1400	0.0855	Isotropic
Var	0.0195	<b>0.0943</b>	0.0983	
MSE	0.0509	0.2343	0.1838	
Bsq	0.0091	<b>0.0027</b>	0.1162	No-bias isotropic
Var	0.2342	0.1133	0.4839	
MSE	0.2433	0.1160	0.6001	
Bsq	0.0284	0.1322	0.0750	Anisotropic
Var	<b>0.0154</b>	0.1046	<b>0.0556</b>	
MSE	0.0437	0.2368	0.1307	
Bsq	<b>0.0023</b>	0.0140	<b>0.0001</b>	No-bias anisotropic
Var	0.0263	0.0967	0.1191	
MSE	<b>0.0286</b>	<b>0.1107</b>	<b>0.1192</b>	

Minimum attained values highlighted for each measurement and synthetic case.

For illustration purposes, Fig. 7a shows the resulting tractography for the noisy cross data set and Fig. 7b after filtering. In Fig. 8, a glyph representation for a region of interest in the axial plane in the Earth experiment is shown. Colors codify orientation. A similar representation for a coronal view in the logarithm experiment is also shown in Fig. 9. In the three cases the anisotropic version of the proposed filtered was applied using the bias correction during  $N = 5$  iterations and  $\lambda = 0.5$ . For the logarithm data set, several scalar measures were determined. In Table 6, the achieved results are shown for the original data, noisy data, isotropic filtering and anisotropic filtering. The filter applied the bias correction with  $N = 5$  iterations and  $\lambda = 0.5$ .

Finally, a last experiment was performed to show the need of the anisotropic filtering. In Fig. 10, a glyph representation with color coding orientation is shown for the original data, the noisy data, isotropic filtering and anisotropic filtering.

### 3.2. Real data

The anisotropic version of the filter with bias correction and for  $N = 5$  and  $\lambda = 0.5$  was applied to the real data. Results are shown in Figs. 11–14. These figures represent axial, sagittal and coronal slices for different measures of anisotropy and orientation. In particular, Fig. 11 shows the FA value, Fig. 12 the planar component, Fig. 13 the spherical component and Fig. 14 a color representation for the main orientation. In each case, the original data set and the result after filtering are shown.

Additionally, in Figs. 15–17, a glyph representation for the regions of interest highlighted in Fig. 14a is shown. In these figures colors codify orientation. As explain in Section 2.5, we have determined nine subregions by thresholding. For the sake of illustration,

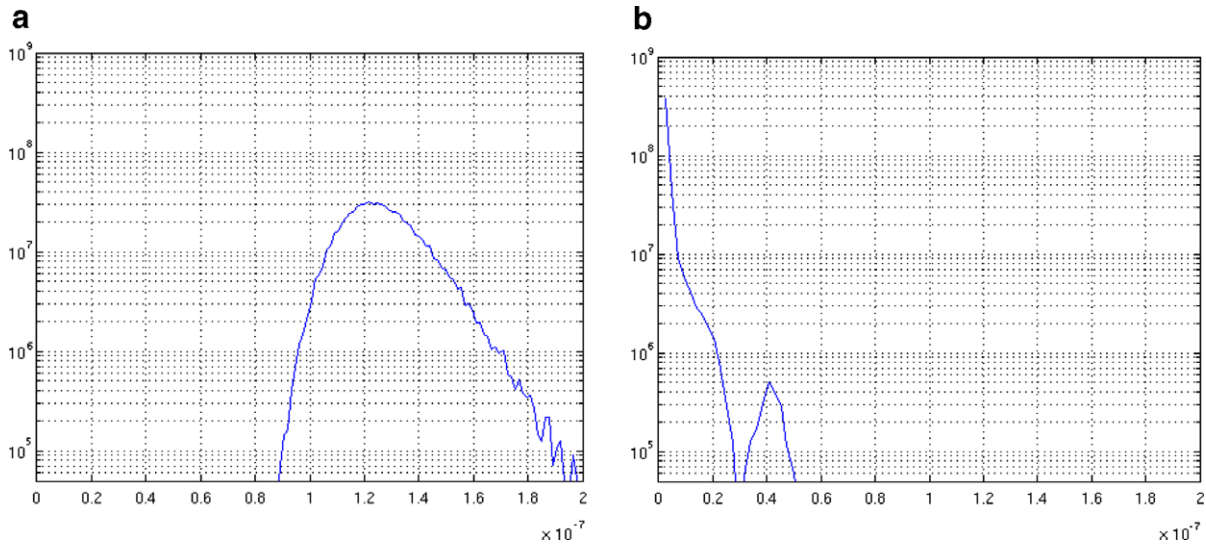


Fig. 5. Experimental distribution for the total variance for the Earth data set: (a) noisy and (b) filtered.

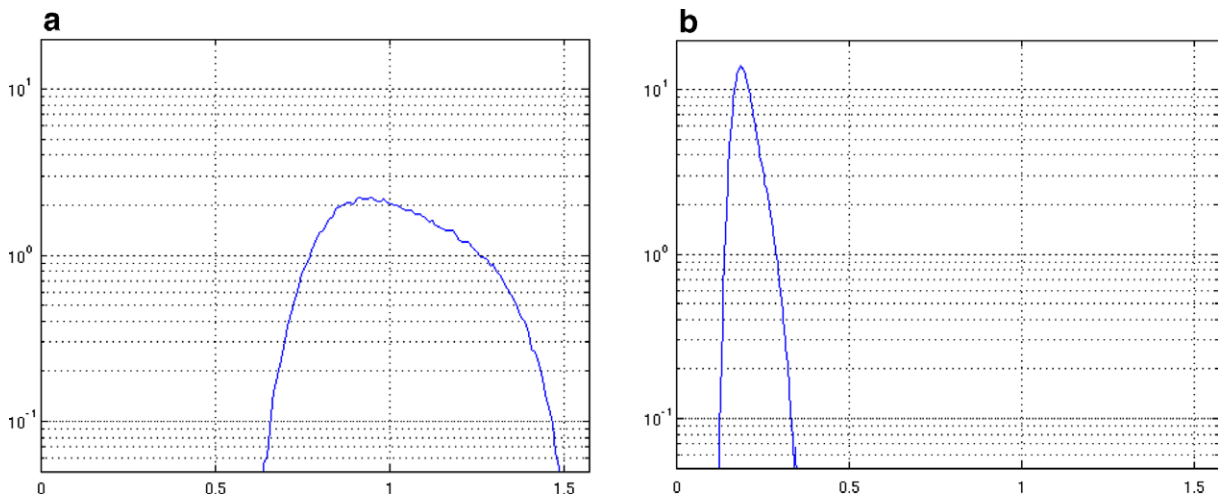


Fig. 6. Experimental distribution for the cone of uncertainty CI for the logarithm data set: (a) noisy and (b) filtered. Angles are measured in radians. The maximum allowed value is  $\pi/2$ .

Table 5

Statistical results for the reconstructed tensor data sets (see the main text for the details)

Synthetic data set	Measure	Noisy/filtered	Mean $\pm$ std	$t$ -Test, $\alpha = 0.05$
Logarithm	$TV \times 10^8$	Noisy	$30.6133 \pm 3.6216$	$p = 0.0$
		Filtered	$1.5564 \pm 0.4809$	
	Cone CI	Noisy	$1.0233 \pm 0.1734$	$p = 0.0$
		Filtered	$0.2027 \pm 0.0343$	
Earth	$TV \times 10^8$	Noisy	$12.6826 \pm 1.4268$	$p = 0.0$
		Filtered	$0.3575 \pm 0.3892$	
	Cone CI	Noisy	$0.5933 \pm 0.0854$	$p = 0.0$
		Filtered	$0.1033 \pm 0.0551$	

in Fig. 18 one of these subregions is shown for the original data set: the blue subregion for the coronal view. We have measured the linear component before and after regularization in these nine subregions of interest. Results are shown in Table 7.

Finally, we have run the fiber stochastic fiber tracking algorithm for the whole brain before and after filtering. We obtained an averaged fiber smoothness  $\bar{\mu} = 147.6 \cdot 10^{-3}$  for the observed data and a value of  $\bar{\mu} = 48.1 \cdot 10^{-3}$  after applying the proposed filtering scheme.

#### 4. Discussion

Synthetic experiments allow us to quantify how the proposed method is able to reduce the amount of noise. Tables 1–4 show some results for the DWI data sets prior to calculating the DTI. These results show how the bias is reduced by the bias correction stage and how the variance is reduced by the Wiener filtering step. In addition, as several versions of the filter were run, comparisons can be made.

Table 2 shows the result after estimating the  $s$  parameter without using the Wiener filtering. It is worth noting that low bias values were obtained at the cost of increasing variances even above the values given in Table 1 for the observed data. The overall result



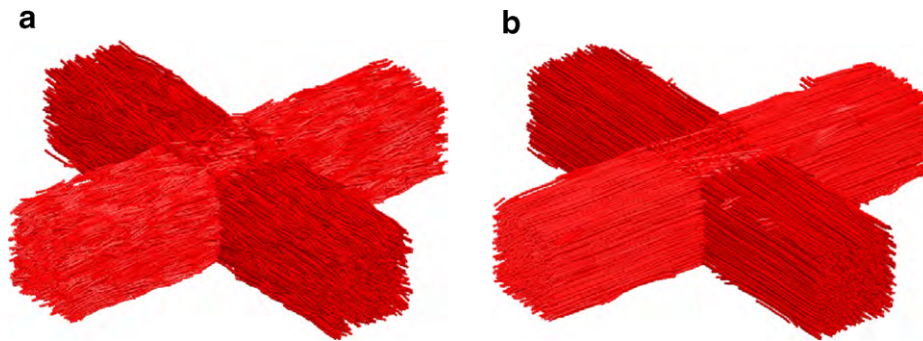


Fig. 7. Tractography results for the cross synthetic image: (a) noisy and (b) filtered for  $\lambda = 0.5$  and  $N = 5$  iterations.

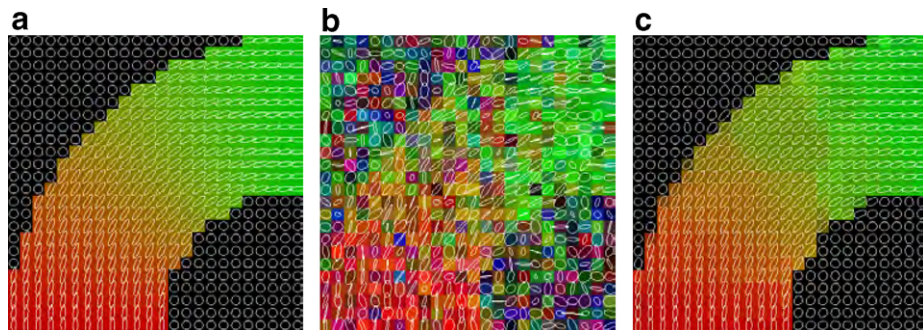


Fig. 8. Glyph representation for the Earth experiment: (a) original, (b) noisy and (c) filtered.

is that the total MSE is very similar to the one given in Table 1. Better results were obtained for the anisotropic estimation than for the isotropic counterpart (bold type fonts were used for the lowest values). It has to be highlighted that the logarithm data set has no discontinuity. That is the reason why the isotropic and anisotropic versions of the filter give rise to very similar results for the variance and MSE (second column in Table 2). However, as in the other two cases (cross and Earth) some discontinuities are present, the anisotropic version of the filter gives lower values for the bias, variance and MSE.

Table 3 provides the achieved results for five iterations under several filter settings. When no bias correction is included, the resulting bias is similar to the one in the observed noisy data, shown in Table 1. The bias correction described in Section 2.3 gives rise to large reductions in the bias which are similar to those shown in Table 2 but with much lower variances, resulting in low MSE values. Comparing isotropic to anisotropic filter types, and generally speaking, the anisotropic version provides lower variance, which leads to lower MSE values as well. This effect can be seen more clearly in Table 4, which gives similar results but for  $N = 10$  iterations. In this latter case, the minimum attained MSE values were obtained for the anisotropic case with bias correction for the three synthetic cases under analysis. The MSE values were 30.30, 35.98 and 13.46 times smaller than the MSE value of the noisy data for the cross, logarithm and Earth cases, respectively, giving rise to an average of 26.6 times smaller MSE value. It is worth noting in the logarithm case, where no discontinuity is present, that the isotropic filter gives very good results and these results are very similar to the anisotropic case for this data set. In the cross experiment, the isotropic filter clearly fails. Though the bias is reduced from 0.0311 to 0.0093 (Table 3) when the bias correction is applied, the variance increases from 0.0302 to 0.2778 (Table 3) giving rise to MSE values of 0.0613 without bias correction and 0.2871 with bias correction. In conclusion, the bias

correction makes the variance increase when the data set has discontinuities. With respect to the Earth case, the situation is even worse: both the bias and the variance increase when the bias correction is included, thus making the MSE value increase from 0.1863 to 0.6697 (Table 3). When the anisotropic version of the filter is considered, more reasonable results are achieved: for the three cases when the bias correction is applied the resulting bias is lower and the variance is of the same order resulting in a minimum MSE value (see Tables 3 and 4).

The synthetic case also allows to evaluate the effect of the proposed method in the corresponding DTI data set. The results for the total variance and the CI of cone of uncertainty, both averaged for  $M = 100$  repetitions, are shown in Figs. 5 and 6 as well as in Table 5. The estimated distribution of the total variance per voxel was highly reduced as shown in Fig. 5 for the Earth with a reduction of the estimated mean of the total variance by 35.47 times on the average as shown in Table 5. In addition, and for the same case, the estimated standard deviation of the total variance was also reduced by a factor of 3.66 on the average. As far as the logarithm case is concerned, the estimated mean of the total variance was reduced by a factor of 19.67 and the estimated standard deviation by a factor of 7.53. We can conclude that the filter makes the total variance of the DTI decrease by an average factor of 27.57 and its standard deviation by an average factor of 5.96. The  $t$ -test was passed with a very low  $p$  value (equal to zero at the numerical precision we have worked with) which leaves no room for doubt about whether the mean value was reduced thanks to the filtering.

Similar results were also obtained for the CI of the cone of uncertainty, as shown in Fig. 6 and Table 5. The estimated mean value for the cone of uncertainty was reduced by an averaged factor of 5.4 and the estimated standard deviation by a factor of 4.1. Similarly the  $t$ -test was also passed with, say, null  $p$ -values as shown in Table 5.

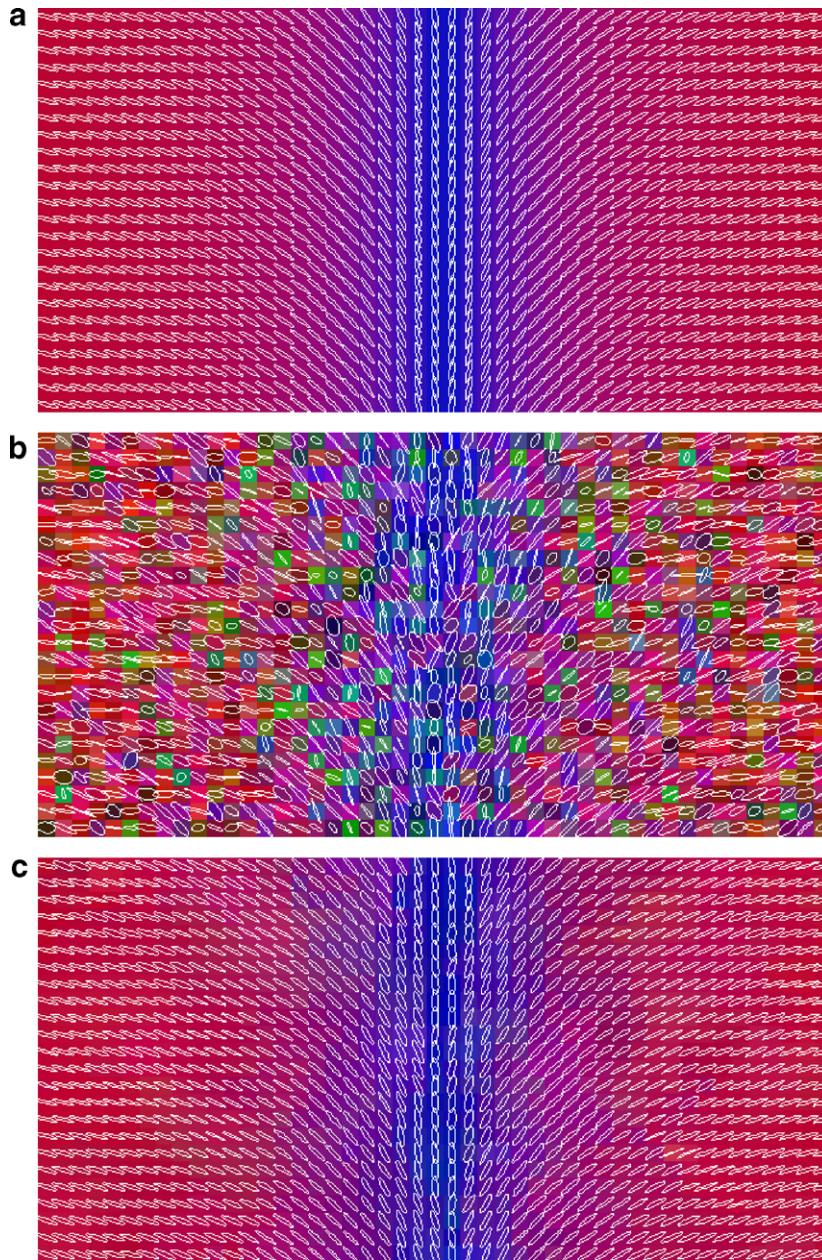


Fig. 9. Glyph representation for the logarithm experiment: (a) original, (b) noisy and (c) filtered.

**Table 6**  
Anisotropy measurements for the logarithm data set

Measure (average)	Original data	Noisy data	Isotropic filtering	Anisotropic filtering
Linear component	0.7142	0.6098	0.7101	0.6933
Planar component	0.1429	0.2673	0.1281	0.1551
Spherical component	0.1429	0.1230	0.1618	0.1516
FA	0.7577	0.7346	0.7441	0.7424
First eigenvalue	7	7.5	6.9	6.8
Second eigenvalue	2	2.7	2.0	2.0
Third eigenvalue	1	0.9	1.1	1.0

In Fig. 7, a representation for the tracts after stochastic fiber tracking is shown before and after filtering. It is clear from Fig. 7b that the tracts are smoother and that the algorithm was able to follow the two paths in the crossing region with less jumps than for the noisy data shown in Fig. 7a. Fig. 8 for the Earth case allows

to see how the anisotropic version of the filter with bias correction is working. The colored glyphs within the ring in Fig. 8c are of the same shape than Fig. 8a: the bias effect of the Rician noise on the tensor shape (shrinking effect) has been corrected. In addition, the orientation is recovered and the regions are not mixed up due to the anisotropic nature of the filter applied, even though the boundaries are not as simple as the ones shown in Fig. 1 and used to estimate the parameters. To further analyze the bias correction performance of the filter in the tensor shapes, we quantify the results achieved for the logarithm experiment. In this case there is no discontinuity and only one region is present, so it is possible to calculate scalar measures and average the results. In Fig. 9, we show the original, noisy and filtered data sets using the same glyph visualization. In this case the view is coronal. Table 6 shows the averaged values for different scalar measures for the original, noisy, isotropic filtered and anisotropic filtered data sets. In both filters the bias correction was applied. In particular, in this table

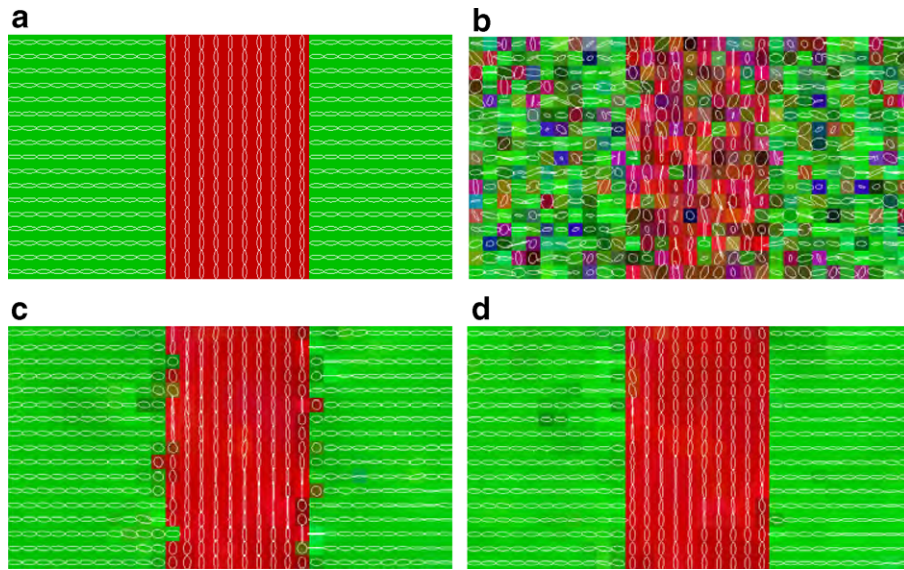


Fig. 10. Glyph representation for the last synthetic experiment: (a) original, (b) noisy, (c) isotropic filtered and (d) anisotropic filtered.

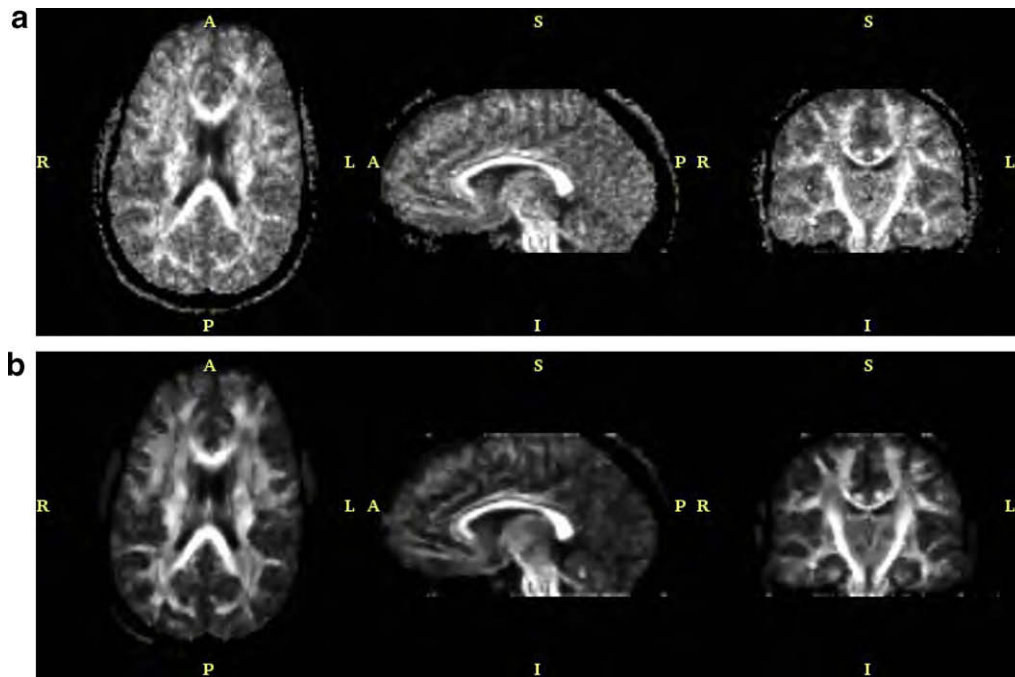


Fig. 11. FA value for the (a) original dataset and (b) after filtering.

we show the linear, planar and spherical components (Westin et al., 2002), the FA value and the three eigenvalues. It is clear that the Rician noise causes a shrinking effect in the tensor shapes by looking at the second column (noisy data) and comparing it with the first one (original data). The linear component has decreased its value from 0.7142 to 0.6098 and the FA value has also decreased its value from 0.7577 to 0.7346. Concerning the eigenvalues, the two larger ones have increased their values and the smaller one has decreased it. The results after filtering clearly shows how these bias effects have been properly corrected giving rise to values very close to the original ones. In this case, the difference between the isotropic and the anisotropic versions of the filter are not very significant as the logarithm data set has no discontinuities.

In the last synthetic experiment, the anisotropic nature of the filter was further analyzed. To better see the results, in this exper-

iment we have increased the amount of Rician noise in the data to levels which usually do not happen in real cases. In Fig. 10c, we can see that the isotropic version of the filter clearly fails close to the boundaries: some of the tensors are wrong and the regions are clearly blurred close to the boundaries. The anisotropic version of the filter shown in Fig. 10d gives a much better result and recovers the main direction nearly perfect in all the cases. This is the reason why, though quantitative measures some times gives worse results for the anisotropic version than for the isotropic<sup>8</sup>, we should always use the anisotropic version in order not to mixed up different regions.

<sup>8</sup> Due to the smaller neighborhoods (oriented as shown in Fig. 1) to estimate the parameters.

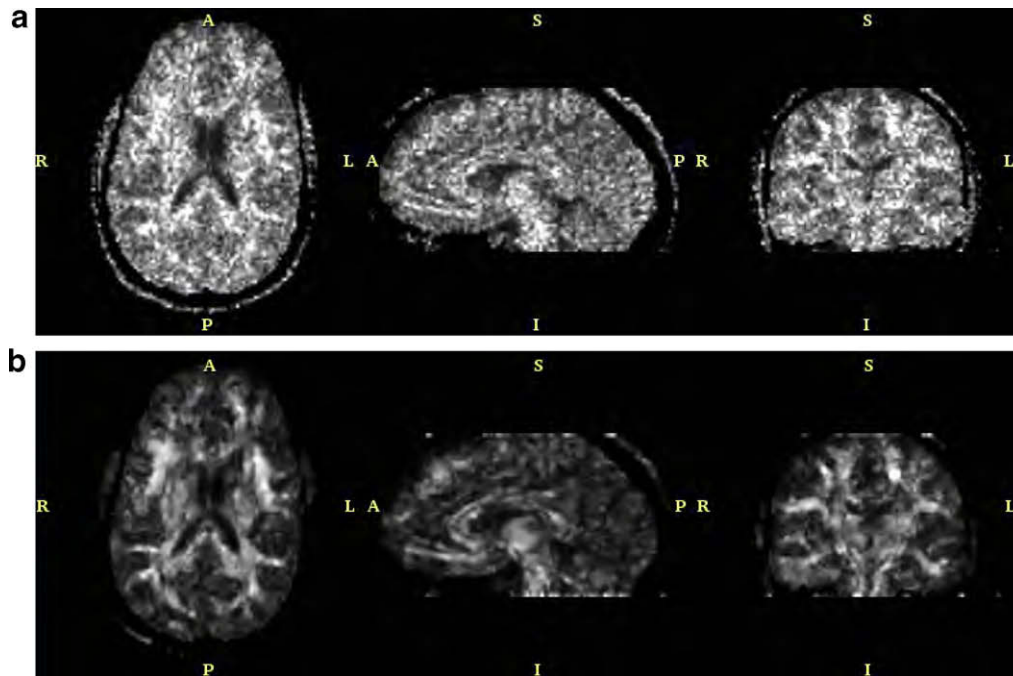


Fig. 12. Planar component for the (a) original dataset and (b) after filtering.

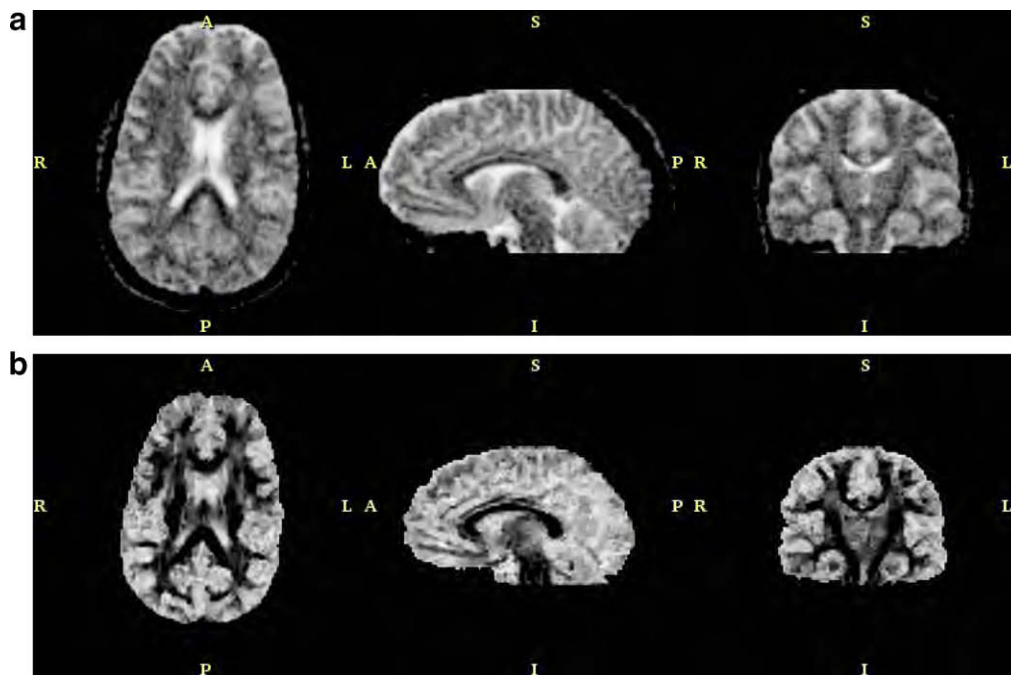


Fig. 13. Spherical component for the (a) original dataset and (b) after filtering.

Concerning the experiment with the real imaging case, Figs. 11–14 show the FA, the planar component, the spherical component and a color-coded representation of the direction of the main eigenvalue, respectively. From these figures it is clear that the amount of uncertainty within the DTI data has been considerably reduced. In particular, the background noise was fairly removed as seen for instance in the FA value in Fig. 11. The uncertainty of the original DTI data has a clearer effect at low FA values (isotropic

areas). Fig. 11b shows a large improvement of the FA value thanks to the variance reduction achieved by the filtering stage. Similar results could have been achieved looking at Figs. 12b and 13b, for the planar and spherical components, respectively. The bias reduction is also an important factor which makes the mentioned figures have larger black areas, i.e., low SNR areas, which were speckled gray for the original noisy data. In our opinion the achieved enhancement for this scalar measures is outstanding. The uncer-

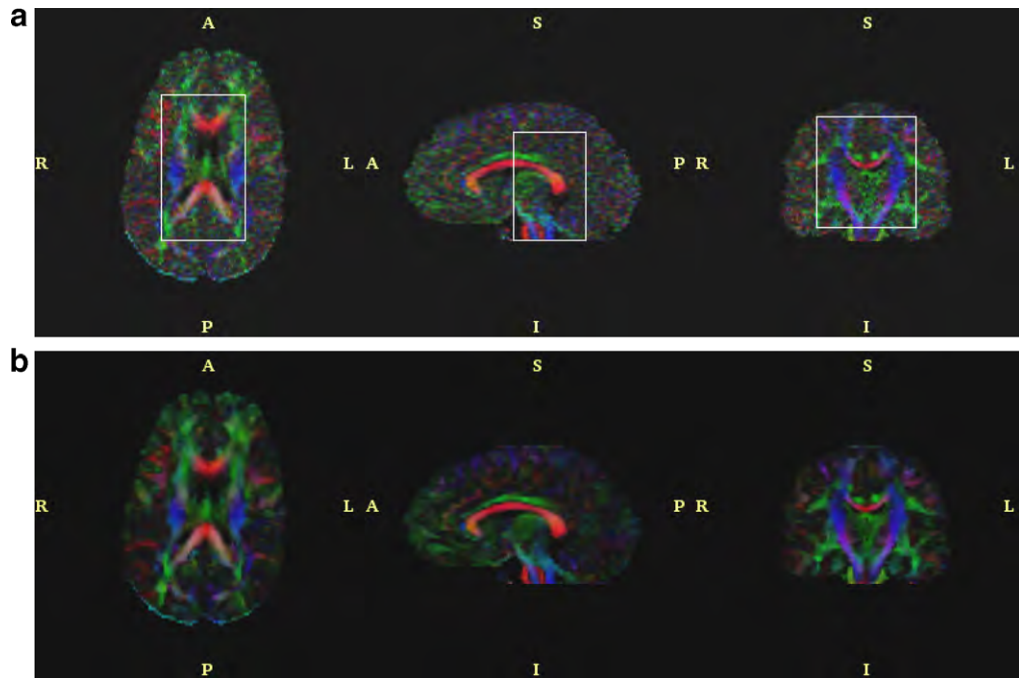


Fig. 14. Color representation of the orientation for the (a) original dataset and (b) after filtering.

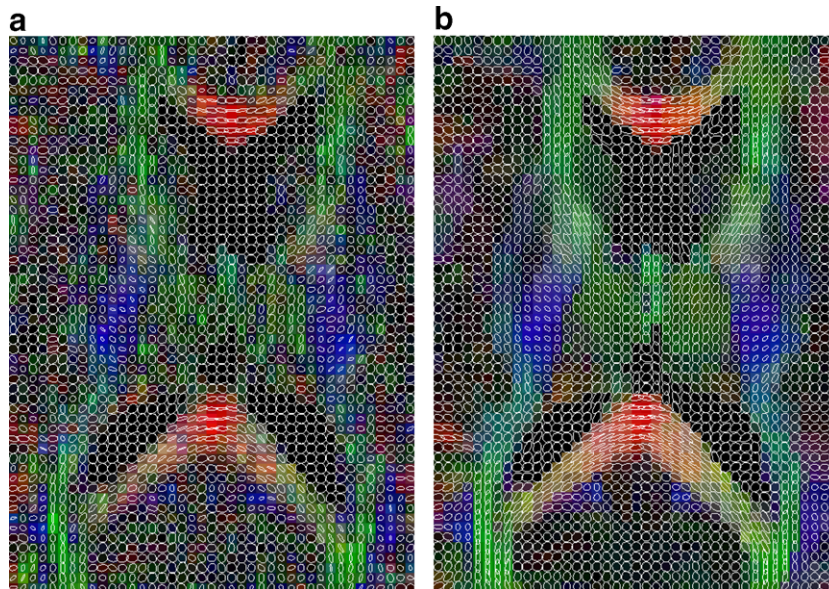


Fig. 15. Glyph representation for an axial region of interest: (a) original and (b) filtered.

tainty for the main diffusivity direction as shown in color in Fig. 14 was also largely reduced. A nicer appearance is achieved for the color image in Fig. 14b when compared to Fig. 14a.

Figs. 15–17, for the three regions of interest shown in Fig. 14a, show in detail how the filtering is working. The view is clearly smoother after filtering and both the shapes and the colors are more regular allowing better interpretation and processing of the DTI. This representation should be interpreted carefully as in each view always one of the colors cannot be represented by the glyphs as it is orthogonal to the ellipse plane. For instance, in Fig. 15 blue regions have circular shape as the diffusion in this case is orthogonal to the image plane. The same happens to red in Fig. 16 and green in Fig. 17. To quantify how the filter is able to cancel out

the bias in the DTI (shrinking effect), in Table 7 we can see the linear component (Westin et al., 2002) in the nine segmented subregions (see Section 3.2 for details about how these regions were determined). These subregions are known to have prolate tensors. In all the nine regions, the averaged linear component has improved its value after filtering. In particular, the linear component averaged in the nine regions before filtering (averaged value of figures in Table 7 top) is 0.4027 and after filtering (averaged value of figures in Table 7 bottom) is 0.6041 which means that the tensors are 50% more linear in those regions after filtering.

Regarding the fiber tracking experiment, the fibers are smoother. To prove this, we can have a look at the achieved results for the fiber smoothness parameter  $\bar{\mu}$  as presented in Section 3.2. These

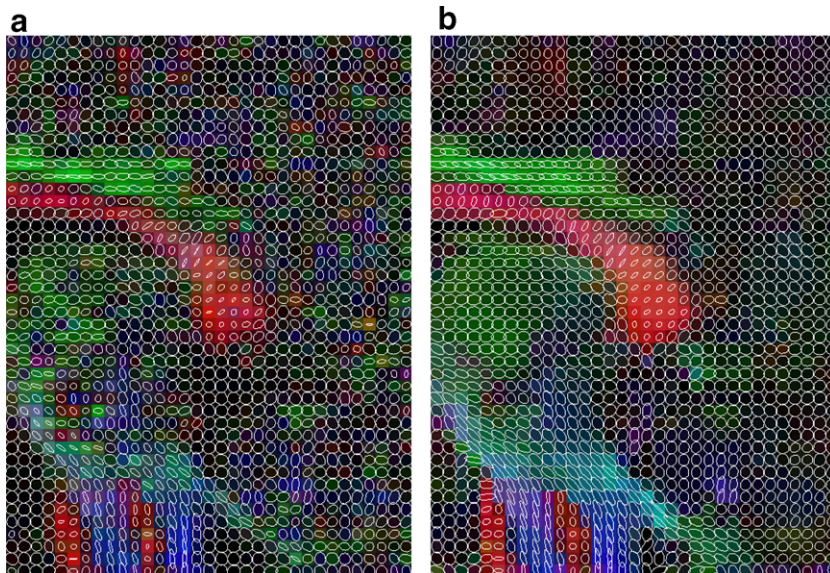


Fig. 16. Glyph representation for an sagittal region of interest: (a) original and (b) filtered.

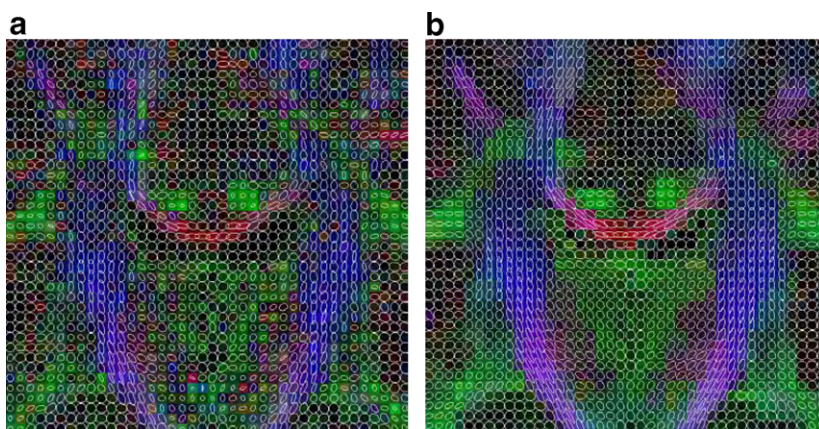


Fig. 17. Glyph representation for an coronal region of interest: (a) original and (b) filtered.



Fig. 18. Selected areas with inferior–superior diffusion (blue areas) in the coronal view.

results mean that after filtering, the obtained fibers were, for the defined smoothness parameter, 3 times smoother on the average.

Concerning the assumptions, as for the first order density function model used, the Rice distribution is well known to be an

Table 7

Averaged linear component in the nine regions of interest

View	Left–right	Anterior–posterior	Inferior–superior
<i>Original data</i>			
Axial	0.5458	0.4333	0.3961
Sagittal	0.4847	0.4502	0.4634
Coronal	0.4893	0.3795	0.4318
<i>After filtering</i>			
Axial	0.7123	0.5751	0.5386
Sagittal	0.6525	0.5932	0.6292
Coronal	0.6815	0.4800	0.5749

appropriate model for the type of data we deal with (see for instance Gudbjartsson and Patz, 1995). As for the second major assumption, i.e., the fact that data are assumed voxelwise uncorrelated, it should be pointed out that in the worst case, that is, in the presence of correlation, our design would only be a suboptimal version of a fully coupled Wiener filter. Therefore, this assumption is far from critical and does not deserve further attention. Additionally, the fact that filter parameters are estimated using data from the surroundings of each voxel under study imposes spatial

coherence, as can be clearly seen from the image experiments included in the paper. The anisotropic estimation was clearly proved in the experiments which avoids blurring the boundaries. The bias correction achieved in the experiments supports the fact that the DTI data should be filtered on the DWI space and not directly on the DTI, otherwise the filter will not be able to correct for the shrinking effect which is not due to the filtering itself but to the noise.

The reader is referred to our previous paper (Martin-Fernandez et al., 2007) to see the effect of the  $\lambda$  parameter on the filtering result. In this paper, this free parameter was analyzed in depth. It concerns mainly to the noise level estimation and the degree of regularization achieved. We have compared the obtained value of the noise covariance using  $\lambda = 0.5$  for the real data set used and the value obtained using a manually segmented region without signal.<sup>9</sup> The corresponding covariance matrices were of the same order of magnitude for the first iteration of the algorithm and also converged to similar values in the subsequent iterations.

With respect to the property of positive-definiteness of the DTI data, we have to highlight that our method does not impose any constraints to the DWI so as to achieved non-negative eigenvalues. For this reason, tensors with negative eigenvalues may happen. However, we have seen that this is not a real problem. After filtering, all the negative eigenvalues happened to occur in the regions outside the patient where no signal is present. No significant impact on the results is thus found. As a result of the filtering process all the negative eigenvalues happening in the brain region were corrected. Imposing this constraint in the model would be a better solution, but it would make the filter more complicated, making its real-time implementation impossible.

The described results have served as a testbed for the presented DWI filtering approach. In summary, it can reduce the amount of noise and uncertainty not only for the original DWI but also for the DTI to levels similar to current state-of-the-art approaches reviewed in Section 1. These approaches are in general computationally very expensive. We here propose an alternative solution which is much more simple albeit giving rise to very acceptable results. The execution times for a compiled version of the filter in an averaged state-of-the-art machine for the real data (whose volume size was  $128 \times 128 \times 54$ ) were 10.3 and 30.6 seconds per iteration for the isotropic and the anisotropic versions, respectively. These figures make the filter useful for real-time applications.

## 5. Concluding remarks

The presented method makes proper use of the available DWI information before determining the tensors. If more gradient directions were provided, the presented method is designed to use the correlation between the signal components of the different DWI images (optimally in the mean square error sense if the parameters were known). The anisotropic approach presented for estimating the model parameters prevents the filter from mixing up data coming from different regions. Additionally, we have included in the filtering process information about the first order density function of the original data, which leads to a decrease in the bias of the filter for non-Gaussian data. Results both on synthetic and on real images support the filtering methodology proposed in the paper.

## Acknowledgements

The authors acknowledge NIH for research Grant P41-RR13218, the Comisión Interministerial de Ciencia y Tecnología, Spain, for research Grant TEC2007-67073/TCM, the Fondo de Investigación

Sanitaria, Spain, for research Grant PI04-1483 and the Junta de Castilla y León for Grants VAO26A07 and VAO27A07.

## Appendix A. Wiener filtering approach

We will assume the model  $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ , where  $\mathbf{X}$  is the ground truth, i.e., the signal without noise, and  $\mathbf{N}$  is the noise. Our goal is to estimate  $\mathbf{X}$  using the observation  $\mathbf{Y}$ . We also assume that the noise  $\mathbf{N}$  has mean  $E\{\mathbf{N}\} = \mathbf{0}$  and covariance matrix  $\mathbf{C}_N = E\{\mathbf{N}\mathbf{N}^T\}$ , which is a symmetric and positive definite matrix. With respect to the signal  $\mathbf{X}$  we will assume a vector of means  $\eta_X = E\{\mathbf{X}\} = E\{\mathbf{Y}\}$  and covariance matrix  $\mathbf{C}_X = E\{\mathbf{X}\mathbf{X}^T\} - \eta_X\eta_X^T$ , which is also symmetric and positive definite. Finally, we will assume that the signal  $\mathbf{X}$  and the noise  $\mathbf{N}$  are uncorrelated so  $E\{\mathbf{X}\mathbf{N}^T\} = \mathbf{0}$ , due to the zero-mean condition of the noise.

The Wiener filter is the optimum linear filter in the mean square error (MSE) sense, provided that the mean is known. Otherwise, more involved approaches are to be applied to guarantee the unbiased condition in the filter output (Ruiz-Alzola et al., 2005). Let  $\mathbf{Z}$  denote the filter output. The linearity of the filter let us write

$$\mathbf{Z} = \mathbf{W}^T \mathbf{Y}, \quad (\text{A.1})$$

where  $\mathbf{W}$  is a coefficient matrix to be determined. The filter output is commonly constrained to be unbiased. This unbiased condition can be easily modeled by modifying Eq. (A.1) to

$$\mathbf{Z} = \mathbf{W}^T [\mathbf{Y} - \eta_X] + \eta_X. \quad (\text{A.2})$$

Needless to say, the mean has to be estimated from data, so this expression in a real-world application is only an approximation. In order to determine the coefficient matrix  $\mathbf{W}$ , the MSE has to be minimized. The solution is well known (Haykin, 1996) and given by

$$\mathbf{Z} = \mathbf{C}_X (\mathbf{C}_X + \mathbf{C}_N)^{-1} [\mathbf{Y} - \eta_X] + \eta_X. \quad (\text{A.3})$$

## Appendix B. On the SNR for Rayleigh and Rice distributed data

Let  $\tilde{Y} = Y_c + jY_s$  be the pre-envelope of the received signal, where  $Y_c$  and  $Y_s$  are the in-phase and quadrature components of the received signal, respectively. Due to noise during the acquisition procedure both components are independent normal variables with the same variance  $\sigma^2$  and means  $s_c$  and  $s_s$ , respectively. This is the same as to assume the model  $\tilde{Y} = \tilde{s} + \tilde{N}$  for pre-envelopes, where  $\tilde{N}$  and  $\tilde{s}$  are the pre-envelopes of the noise and noise-free signal components, respectively. The noise pre-envelope is  $\tilde{N} = N_c + jN_s$ , where  $N_c$  and  $N_s$  are the in-phase and quadrature components of the noise, respectively. Both components are independent normal variables with same variance  $\sigma^2$  and zero means, respectively. The noise natural envelope is given by

$$N = |\tilde{N}| = \sqrt{N_c^2 + N_s^2}. \quad (\text{B.1})$$

It can be easily shown that this variable is Rayleigh-distributed and its probability density function is given by

$$P_N(n) = \frac{n}{\sigma^2} \exp\left(-\frac{n^2}{2\sigma^2}\right) u(n), \quad (\text{B.2})$$

where  $u(\cdot)$  is the step function. Its mean and mean-squared value are given by

$$E\{N\} = \sigma \sqrt{\frac{\pi}{2}} \quad \text{and} \quad E\{N^2\} = 2\sigma^2. \quad (\text{B.3})$$

Let define the *signal-to-noise ratio* (SNR) for a variable  $X$  as

$$\text{SNR} = \frac{E\{X\}}{\sqrt{E\{X^2\} - E^2\{X\}}}. \quad (\text{B.4})$$

<sup>9</sup> Regions included in the DWI outside the patient.

Thus, it can be shown that for a Rayleigh variable the SNR is the constant

$$\text{SNR} = \sqrt{\frac{\pi}{4-\pi}} \approx 1.913. \quad (\text{B.5})$$

In the absence of noise, the pre-envelope  $\bar{s} = s_c + js_s$  of the received signal has a natural envelope given by

$$s = |\bar{s}| = \sqrt{s_c^2 + s_s^2}. \quad (\text{B.6})$$

The goal is to estimate this natural envelope  $s$  out of the natural envelope of the received signal  $\tilde{Y}$  that can be written as

$$Y = |\tilde{Y}| = \sqrt{Y_c^2 + Y_s^2}. \quad (\text{B.7})$$

In this case, this variable is Rician distributed and its probability density function is given by

$$P_Y(y) = \frac{y}{\sigma^2} \exp\left(-\frac{y^2 + s^2}{2\sigma^2}\right) I_0\left(\frac{ys}{\sigma^2}\right) u(y), \quad (\text{B.8})$$

where  $I_n(\cdot)$  is the modified Bessel function of first kind and order  $n$ . It is worth mentioning the dependence of this distribution on the noise-free component  $s$ . Inference on this component  $s$  out the observed data  $Y$  is possible in a closed manner as we will show next.

The mean of this distribution is given by

$$E\{Y\} = \sigma \sqrt{\frac{\pi}{2}} \exp\left(-\frac{s^2}{4\sigma^2}\right) \left[ \left(1 + \frac{s^2}{2\sigma^2}\right) I_0\left(\frac{s^2}{4\sigma^2}\right) + \frac{s^2}{2\sigma^2} I_1\left(\frac{s^2}{4\sigma^2}\right) \right] \quad (\text{B.9})$$

and its mean-squared value

$$E\{Y^2\} = 2\sigma^2 + s^2. \quad (\text{B.10})$$

By using Eqs. (B.4), (B.9) and (B.10) the SNR for a Rice variable can be written as

$$\begin{aligned} \text{SNR} &= B(\gamma) \\ &= \frac{\sqrt{\frac{\pi}{2}} \exp\left(-\frac{\gamma^2}{4}\right) \left[ \left(1 + \frac{\gamma^2}{2}\right) I_0\left(\frac{\gamma^2}{4}\right) + \frac{\gamma^2}{2} I_1\left(\frac{\gamma^2}{4}\right) \right]}{\sqrt{2 + \gamma^2 - \frac{\pi}{2} \exp\left(-\frac{\gamma^2}{4}\right) \left[ \left(1 + \frac{\gamma^2}{2}\right) I_0\left(\frac{\gamma^2}{4}\right) + \frac{\gamma^2}{2} I_1\left(\frac{\gamma^2}{4}\right) \right]^2}}, \end{aligned} \quad (\text{B.11})$$

where  $\gamma = s/\sigma$  is known as the *coherent to non-coherent* ratio. In Eq. (B.11) the dependence of the SNR on the  $\gamma$  parameter was made explicit by the use of function  $B(\cdot)$ . It is worth noting that the SNR has only one degree of freedom as the dependence on the parameters  $s$  and  $\sigma$  is exclusively through the parameter  $\gamma$ . We have the following limiting values for this SNR

$$\lim_{\gamma \rightarrow 0} \text{SNR} = \lim_{\gamma \rightarrow 0} B(\gamma) = \sqrt{\frac{\pi}{4-\pi}} \approx 1.913, \quad (\text{B.12})$$

$$\lim_{\gamma \rightarrow \infty} \text{SNR} = \lim_{\gamma \rightarrow \infty} B(\gamma) = \gamma. \quad (\text{B.13})$$

Eq. (B.12) corresponds to the low SNR case (Rayleigh distribution) and Eq. (B.13) to the high SNR case (normal distribution). The bias effect has to be considered mainly for low SNR. For high SNR as the Rice distribution approaches a Gaussian distribution, the Wiener solution is the general optimum solution with no bias. In Fig. 2 we can see the SNR as a function of the  $\gamma$  parameter.

## Appendix C. Synthetic tensor data generation

### C.1. Earth

Consider the parametric family  $(x, y, z) = (\rho \cos(\theta), \rho \sin(\theta), \alpha)$ , where  $\theta \in (-\pi, \pi)$  is the curve parameter and  $\rho > 0$  and  $\alpha$  are con-

stant for each curve. Setting  $\rho$  and  $\alpha$  and varying  $\theta$  a curve is generated. The tangent vector for this curve is given by the first derivative of the curve with respect to the free parameter  $\theta$  as

$$\left( \frac{\partial x}{\partial \theta}, \frac{\partial y}{\partial \theta}, \frac{\partial z}{\partial \theta} \right) = (-\rho \sin(\theta), \rho \cos(\theta), 0) = (-y, x, 0). \quad (\text{C.1})$$

The unitary eigenvectors for the tensor field are given by

$$\begin{cases} \mathbf{v}_1 = \left( \frac{-y}{\sqrt{x^2+y^2}}, \frac{x}{\sqrt{x^2+y^2}}, 0 \right)^T, \\ \mathbf{v}_2 = \left( \frac{x}{\sqrt{x^2+y^2+1}}, \frac{y}{\sqrt{x^2+y^2+1}}, \frac{1}{\sqrt{x^2+y^2+1}} \right)^T, \\ \mathbf{v}_3 = \left( \frac{-x}{\sqrt{x^4+y^4+x^2y^2+x^2+y^2}}, \frac{-y}{\sqrt{x^4+y^4+x^2y^2+x^2+y^2}}, \frac{x^2+y^2}{\sqrt{x^4+y^4+x^2y^2+x^2+y^2}} \right)^T, \end{cases} \quad (\text{C.2})$$

where  $\mathbf{v}_1$  follows the tangent field given by Eq. (C.1). The other two eigenvectors are chosen so as to form an orthonormal set.

### C.2. Logarithm

The parametric family for the curve is  $(x, y, z) = (\rho \cos(\theta), \rho \sin(\theta), \ln \rho + \alpha)$ , where  $\rho > 0$  is the curve parameter and  $\theta \in (-\pi, \pi)$  and  $\alpha$  are constant for each curve. Setting  $\theta$  and  $\alpha$  and varying  $\rho$  a curve is generated. The tangent vector for this curve is given by the first derivative of the curve with respect to the free parameter  $\rho$  as

$$\left( \frac{\partial x}{\partial \rho}, \frac{\partial y}{\partial \rho}, \frac{\partial z}{\partial \rho} \right) = \left( \cos(\theta), \sin(\theta), \frac{1}{\rho} \right) = \left( \frac{x}{\rho}, \frac{y}{\rho}, \frac{1}{\rho} \right). \quad (\text{C.3})$$

The unitary eigenvectors for the tensor field are given by

$$\begin{cases} \mathbf{v}_1 = \left( \frac{x}{\sqrt{x^2+y^2+1}}, \frac{y}{\sqrt{x^2+y^2+1}}, \frac{1}{\sqrt{x^2+y^2+1}} \right)^T, \\ \mathbf{v}_2 = \left( \frac{-y}{\sqrt{x^2+y^2}}, \frac{x}{\sqrt{x^2+y^2}}, 0 \right)^T, \\ \mathbf{v}_3 = \left( \frac{-x}{\sqrt{x^4+y^4+x^2y^2+x^2+y^2}}, \frac{-y}{\sqrt{x^4+y^4+x^2y^2+x^2+y^2}}, \frac{x^2+y^2}{\sqrt{x^4+y^4+x^2y^2+x^2+y^2}} \right)^T, \end{cases} \quad (\text{C.4})$$

where  $\mathbf{v}_1$  follows the tangent field given by Eq. (C.3). The other two eigenvectors are chosen so as to form an orthonormal set.

### C.3. Cross

In this case the eigenvectors are defined by

$$\begin{cases} \mathbf{v}_1 = \begin{cases} (1, 0, 0)^T & \text{if } |y| < \alpha/2, \\ (0, 1, 0)^T & \text{otherwise,} \end{cases} \\ \mathbf{v}_2 = \begin{cases} (0, 1, 0)^T & \text{if } |y| < \alpha/2, \\ (1, 0, 0)^T & \text{otherwise,} \end{cases} \\ \mathbf{v}_3 = (0, 0, 1)^T, \end{cases} \quad (\text{C.5})$$

where  $\alpha$  is a constant representing the size of the crossing region. The corresponding eigenvalues are given by

$$\lambda_1 = \begin{cases} 7 & \text{if } |y| < \alpha/2, \quad |z| < \gamma/2, \\ 7 & \text{if } |x| < \alpha/2, \quad |y| > \alpha/2, \quad |z| < \gamma/2, \\ 1 & \text{otherwise,} \end{cases} \quad (\text{C.6})$$

$$\lambda_2 = \begin{cases} 7 & \text{if } |x| < \alpha/2, \quad |y| < \alpha/2, \quad |z| < \gamma/2, \\ 2 & \text{if } |y| < \alpha/2, \quad |x| > \alpha/2, \quad |z| < \gamma/2, \\ 2 & \text{if } |y| > \alpha/2, \quad |x| < \alpha/2, \quad |z| < \gamma/2, \\ 1 & \text{otherwise,} \end{cases} \quad (\text{C.7})$$

$$\lambda_3 = 1,$$



where  $\gamma$  is a constant representing the height (along the z component) of the cross.

#### Appendix D. Smoothness along fibers

A fiber  $\mathbf{F}$  is given by a sorted collection of  $K$  points

$$\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_k, \dots, \mathbf{F}_K)^T, \quad (\text{D.1})$$

where each  $\mathbf{F}_k = (x_k, y_k, z_k)$  is a point in the Euclidean 3D space. A collection of tangent vectors to the fiber  $\mathbf{F}$  can be defined by taking the first difference  $\mathbf{dF}$  as

$$\mathbf{dF} = (\mathbf{F}_2 - \mathbf{F}_1, \dots, \mathbf{F}_{k+1} - \mathbf{F}_k, \dots, \mathbf{F}_K - \mathbf{F}_{K-1})^T, \quad (\text{D.2})$$

where  $\mathbf{dF}$  has  $K - 1$  elements. The angle  $\theta$  between two given vectors  $\mathbf{a}$  and  $\mathbf{b}$  is known to be

$$\theta = \arccos \left( \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|} \right), \quad (\text{D.3})$$

where  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$  stand for the scalar product and the norm, respectively. Using matrix operations we can determine the angles  $\theta$

$$\theta = (\theta_1, \dots, \theta_k, \dots, \theta_{K-2})^T, \quad (\text{D.4})$$

spanned by adjacent tangent vectors to the fiber as

$$\theta_k = \arccos \left( \frac{(\mathbf{dF}_1 \mathbf{dF}_2^T)_{kk}}{\sqrt{(\mathbf{dF}_1 \mathbf{dF}_1^T)_{kk} (\mathbf{dF}_2 \mathbf{dF}_2^T)_{kk}}} \right) \quad (\text{D.5})$$

for  $1 \leq k \leq K - 2$ , where  $(\cdot)_{ij}$  stands for the element at position  $(i, j)$  and  $\mathbf{dF}_1$  and  $\mathbf{dF}_2$  are the first  $K - 2$  and the last  $K - 2$  elements of  $\mathbf{dF}$  respectively.

We define the smoothness  $\mu$  for a fiber by

$$\mu = \frac{1}{K-2} \sum_{k=1}^{K-2} (1 - |\cos(\theta_k)|), \quad (\text{D.6})$$

with  $0 \leq \mu \leq 1$ . Parallel vectors give rise to a  $\cos(\theta_k)$  value of 1, anti-parallel vectors a value of  $-1$  and orthogonal vectors a value of 0. Consequently, the lower the  $\mu$  value, the smoother the fiber.

#### References

- Basser, P.J., 1997. Quantifying errors in fibre-tract direction and diffusion tensor field maps resulting from MR noise. In: Proceedings of the Fifth Annual Meeting of the International Society for Magnetic Resonance in Medicine, ISMRM, Vancouver, BC, Canada, April 1997, p. 1740.
- Basser, P.J., Pajevic, S., 2000. Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. *Magnetic Resonance in Medicine* 44, 41–50.
- Basser, P.J., Mattiello, J., LeBihan, D., 1994. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B* 103, 247–254.
- Basu, S., Flechter, T., Whitaker, R., 2006. Rician noise removal in diffusion tensor MRI. *Lecture Notes in Computer Science* 4190, 117–125.
- Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of Royal Statistical Society, Series B* 36, 192–236.
- Chefd'hotel, C., Tschumperle, D., Deriche, R., Faugeras, O., 2002. Constrained flows of matrix-valued functions: application to diffusion tensor regularization. *Lecture Notes in Computer Science* 2350, 251–265.
- Chen, B., Hsu, E.W., 2005. Noise removal in magnetic resonance diffusion tensor imaging. *Magnetic Resonance in Medicine* 54, 393–407.
- Coulon, O., Alexander, D.C., Arridge, S., 2004. Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis* 8, 47–67.
- Fillard, P., Pennec, X., Arsigny, V., Ayache, N., 2007. Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics. *IEEE Transactions on Medical Imaging* 26, 1472–1482.
- Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine* 34, 910–914.
- Hagmann, P., Thiran, J., Jonasson, L., Vandergheynst, P., Clarke, S., Maeder, P., Meuli, R., 2003. DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *NeuroImage* 19, 545–554.
- Haykin, S., 1996. *Adaptive Filter Theory*, third ed. Prentice Hall, Englewood Cliffs, NJ, USA.
- Hillery, A.D., Chin, R.T., 1991. Iterative Wiener filters for image restoration. *IEEE Transactions on Signal Processing* 39, 1892–1899.
- Jones, D.K., 2003. Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. *Magnetic Resonance in Medicine* 49, 7–12.
- Jones, D.K., Basser, P.J., 2004. Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine* 52, 979–993.
- Martin-Fernandez, M., Westin, C.-F., Alberola-Lopez, C., 2004. 3D Bayesian regularization of diffusion tensor MRI using multivariate Gaussian Markov random fields. *Lecture Notes in Computer Science* 3216, 351–359.
- Martin-Fernandez, M., Alberola-Lopez, C., Ruiz-Alzola, J., Westin, C.-F., 2007. Sequential anisotropic Wiener filtering applied to 3D MRI data. *Magnetic Resonance Imaging* 25, 278–292.
- McGraw, T., Vemuri, B.C., Chen, Y., Mareci, T., Rao, M., 2004. DT-MRI denoising and neural fiber tracking. *Medical Image Analysis* 8, 95–111.
- Parker, G.J., Schnabel, J.A., Symms, M.R., Werring, D.J., Barker, G.J., 2000. Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging. *Journal of Magnetic Resonance Imaging* 11, 702–710.
- Pennec, X., Fillard, P., Ayache, N., 2006. A Riemannian framework for tensor computing. *International Journal of Computer Vision* 66, 41–66.
- Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 629–639.
- Poupon, C., Clark, C.A., Frouin, V., Regis, J., Le Bihan, D., Bloch, I., Mangin, J.F., 2000. Regularization diffusion-based direction maps for the tracking of brain white matter fascicles. *NeuroImage* 12, 184–195.
- Ruiz-Alzola, J., Alberola-Lopez, C., Westin, C.F., 2005. Kriging filters for multidimensional signal processing. *Signal Processing* 85, 413–439.
- Sijbers, J., den Dekker, A.J., 2004. Maximum likelihood estimation of signal amplitude and noise variance from MR data. *Magnetic Resonance in Medicine* 51, 586–594.
- Stejskal, E.O., Tanner, J.E., 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics* 42, 288–292.
- Tschumperle, D., Deriche, R., 2003. DT-MRI images: estimation, regularization and application. *Lecture Notes in Computer Science* 2809, 530–541.
- Wang, Z., Vemuri, B.C., Chen, Y., H Mareci, T., 2004. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. *IEEE Transactions on Medical Imaging* 23, 930–939.
- Westin, C.-F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R., 2002. Processing and visualization for diffusion tensor MRI. *Medical Image Analysis* 6, 93–108.