Clinical DT-MRI Estimation, Smoothing and Fiber Tracking with Log-Euclidean Metrics

Pierre Fillard*, Xavier Pennec*, Vincent Arsigny*, and Nicholas Ayache*
* Asclepios Reseach Team, INRIA Sophia Antipolis,
2004 route des Lucioles, BP 93, 06902 Sophia Antipolis, France.

***Abstract*— Diffusion tensor MRI (DT-MRI or DTI) is an imaging modality that is gaining importance in clinical applications. However, in a clinical environment, data have to be acquired rapidly, often at the expense of the image quality. This often results in DTI datasets that are not suitable for complex post-processing like fiber tracking. We propose a new variational framework to improve the estimation of DT-MRI in this clinical context. Most of the existing estimation methods rely on a log-Gaussian noise (Gaussian noise on the image logarithms), or a Gaussian noise, that do not reflect the Rician nature of the noise in MR images with low SNR. With these methods, the Rician noise induces a shrinking effect: the tensor volume is underestimated when other noise models are used for the estimation. In this paper, we propose a maximum likelihood strategy that fully exploits the assumption of a Rician noise. To further reduce the influence of the noise, we optimally exploit the spatial correlation by coupling the estimation with an anisotropic prior previously proposed on the spatial regularity of the tensor field itself, which results in a maximum a posteriori estimation. Optimizing such a non-linear criterion requires adapted tools for tensor computing. We show that Riemannian metrics for tensors, and more specifically the Log-Euclidean metrics, are a good candidate and that this criterion can be efficiently optimized. Experiments on synthetic data show that our method correctly handles the shrinking effect even with very low SNR, and that the positive definiteness of tensors is always insured. Results on real clinical data demonstrate the truthfulness of the proposed approach and show promising improvements of fiber tracking in the brain and the spinal cord.**

***Index Terms*— DT-MRI, estimation, smoothing, tractography, Log-Euclidean**

I. Introduction

Diffusion tensor MRI (DT-MRI or DTI) [1] is a unique tool to assess in vivo oriented structures within tissues via the measure of water diffusion. However, using such an imaging modality in a clinical environment is difficult and acquisitions generally have a limited number of encoding gradients and low signal-to-noise ratios (SNR). Indeed, pathologies often prevent the patient from staying too long in the same position in the scanner. This short scanning time prevents from acquiring and averaging the large number of gradient directions that is necessary for enhancing the SNR. Moreover, the devices that are usually available for clinical purposes (at least in France) usually offer only low quality DWI datasets (generally 6 gradient directions with 4 repeated scans). It is known that the estimation of the diffusion tensor field from diffusion weighted images (DWI) is noise-sensitive. Consequently, clinical DTI is very often not suitable for complex post processing, like fiber tracking. For these reasons, there has been a growing interest in the regularization of tensor images. In the following we quickly summarize the state of the art in diffusion tensor estimation and regularization. Available methods generally perform each of theses two steps independently. We propose in this paper to couple them in a single Maximum-a-Posteriori (MAP) estimation that better captures the information in these intrinsically noisy clinical images. Note that a preliminary version of this work was previously presented in [2].

The Stejskal-Tanner diffusion equation [1] relates the diffusion tensor $\mathbf{D}$ to each DWI with $S_i = S_0 \exp(-b\mathbf{g_i}^\top \mathbf{D}\mathbf{g_i})$, where S_i is the original DWI corresponding to the encoding gradient $\mathbf{g_i}$, S_0 an image with a null gradient, and b the diffusion factor. To get a linear system, one usually takes the logarithm of the DWI [3]–[5]. Solving the linearized system in a least square sense leads to the minimization of a quadratic criterion with algebraic methods. However, doing this corresponds to a maximum likelihood (ML) estimation with a Gaussian noise on the logarithm of the images (we call this noise model log-Gaussian): $\log(\hat{S}_i) = \log(S_0) - b\mathbf{g_i}^\top \mathbf{D}\mathbf{g_i} + N_i(0, \sigma_{lg})$, where $\hat{S}_i$ is the measured DWI intensity, and $N_i(0, \sigma_{lg})$ is a centered Gaussian noise of variance σ_{lg}. One can wonder if this assumption correctly reflects the noise appearing in real DW images. In fact, when the SNR is high, one can show that the noise on the image logarithms is well approximated by a Gaussian distribution [6], which justifies the linearization of the diffusion equation. In the same conditions, the noise can also be well approximated by a Gaussian distribution within the brain [7]: $\hat{S}_i = S_i + N_i(0, \sigma_g)$.

However, for low SNR images typical of clinical acquisitions, the real nature of the noise is Rician, which corresponds to taking the magnitude of a complex signal whose real and imaginary parts are corrupted by a Gaussian noise [8]: $\hat{\mathbf{S}}_\mathbf{i} = \sqrt{[\mathbf{S_i} + \mathbf{N_1}(\mathbf{0}, \sigma_\mathbf{r})]^\mathbf{2} + \mathbf{N_2}(\mathbf{0}, \sigma_\mathbf{r})^\mathbf{2}}$, where N_1 and N_2 are two independent Gaussian noises. This is equivalent to adding Gaussian noise on the k-space before computing the signal magnitude. Wang et al. proposed for instance an estimation criterion on the complex DWI signal that is adapted to this type of noise [9]. We propose in this paper a ML strategy which fully exploits the a priori knowledge on the probability density function (pdf) of the Rician noise, and which does not need to access the full complex signal but only its magnitude, i.e., the standard DWI intensity.

In view of fiber reconstruction, the diffusion tensor field

needs to be regularized without blurring the transitions between distinct fiber tracts, which delimit anatomical and functional brain regions. A first idea consists in smoothing independently each DWI, as done for instance very recently in [10] with a Rician noise model. This results in a smoother tensor field that conserves some of the transitions. However, it also blurs the transitions between some regions as all the boundaries of the tensor field are not visible in each DWI taken separately. For instance, in brain DTI, only the combination of all the DW images reveals the complex neural structure of the white matter. Consequently, we believe that it is better to detect the transitions on the tensor field itself.

Some regularization methods have been proposed as a post-processing step after the estimation of the tensor field. For instance, [11] regularizes the principal eigenvector (associated to the largest eigenvalue), while [12] uses the spectral decomposition of tensors to independently regularize their eigenvectors and eigenvalues. These two methods rely on the spectral decomposition of tensors: considering only the principal eigenvector for smoothing induces a loss of information and there is an uncertainty of the spectral decomposition in regions with flat tensors. Moreover, discontinuity problems arise when smoothing the field of orthogonal matrices as the spectral decomposition is not unique.

By contrast, it would be more interesting to consider the regularization as a spatial prior on the tensor field during the estimation step itself. This would allow to optimally weight the information brought by the observed images and the expected spatial regularity. Such a MAP estimation should extract the maximal amount of meaningful information from very noisy clinical DWI data. In that spirit, [9] proposed to parameterize the space of tensors by the vector space of lower triangular matrices thanks to the Cholesky decomposition. A lower triangular matrix M is such that $\forall (i,j)$, if $i > j$ then $M(i,j) = 0$ (all terms above the diagonal are null). This provided them with a computational framework that could handle a joint estimation and regularization of the tensor field from the complex DWI signal.

However, it is quite difficult to understand the structure of the tensor space given by the Cholesky parameterization. We believe that the use of Riemannian metrics as in [13] is a better theoretical choice. Among these metrics, the Log-Euclidean (LE) ones [14] also turn tensors into a vector space. Section II investigates and compares these different tools for computing with tensors. It appears that LE metrics are computationally as efficient as the Cholesky decomposition while ensuring the positive definiteness of tensors and canceling the swelling effect in regularization.

Based on the LE computational framework, we detail in Section III the variational method of the joint estimation and smoothing of DTI. We first derive the ML estimation of tensors with the log-Gaussian, Gaussian and Rician noise models. We show that the Rician noise induces a *shrinking effect* when other noise models are used for tensor estimation. By adding an anisotropic spatial prior in a second step, we turn these three ML into three MAP methods where the estimation and the regularization of the tensor field are jointly performed.

A quantitative analysis of our 6 new methods on synthetic data shows in Sec. IV that the Rician ML method correctly handles the shrinking effect even with a very low SNR, while other methods under-estimate the tensor volume by as much as 40%. Then, we switch to experiments with real clinical data on a medium quality brain DTI dataset (tumor case with only 6 gradient directions) and a low quality experimental acquisition of the spinal cord (same sequence). The visual inspection of the reconstructed tensor field shows that the MAP method nicely preserves the separation between different regions. A careful analysis in two specific regions shows that ML and MAP Rician methods exhibit a larger tensor volume and ADC than Gaussian and log-Gaussian methods. Last but not least, we illustrate that the MAP Rician method qualitatively improves the fiber tracking.

II. Tools for Tensor Computing

Tensor computing is difficult due to the severe limitations of the standard Euclidean calculus: while convex operations are stable (e.g. the mean of a set of tensors is a tensor), one can quickly reach the boundaries of the space with complex operations (e.g., gradient descent or partial differential equations (PDE)) and null or negative eigenvalues may appear. To overcome this limitation, [9] proposed to parameterize a tensor $\mathbf{D}$ by its Cholesky factor. A Cholesky decomposition of $\mathbf{D}$ is given by: $\mathbf{D} = \mathbf{L}\mathbf{L}^\top$, where $\mathbf{L}$ is a lower triangular matrix. For any lower triangular matrix $\mathbf{L}$, the matrix $\mathbf{L}\mathbf{L}^\top$ is positive (this is easily shown using a SVD decomposition). However, the definiteness is not ensured (null eigenvalues are possible). The authors argue that forbidding negative eigenvalues is sufficient because one cannot numerically distinguish very small eigenvalues from null ones. Thus, forbidding explicitly null eigenvalues has no practical justification.

While we agree that very small eigenvalues are not distinguishable from null ones, we believe that both are very unlikely to exist from a physical point of view, and should be as far as possible from any reference tensor. In other words, a tensor with very small eigenvalues has a very low probability to appear, as well as a tensor with very large eigenvalues. Both of them must be numerically nearly impossible to reach.

The recently proposed Riemannian metrics offer a solution to this constraint: In [13], [15]–[17] an affine-invariant Riemannian metric is proposed and the tensor space is replaced by a regular manifold where matrices with null and negative eigenvalues are at an infinite distance from any tensor. However, computations with this metric are time-consuming since they extensively use the matrix exponential, logarithm, square root and inverse. A novel family of Riemannian metrics, called Log-Euclidean, combines the properties of the affine-invariant family with a computational cost similar to the Euclidean case [14]. The basic idea is to take the matrix logarithm of a tensor $\mathbf{D}$: $\mathbf{L} = \log(\mathbf{D})$, and to run computations on $\mathbf{L}$. The new processed value $\tilde{\mathbf{L}}$ obtained is turned back into a tensor by taking the matrix exponential: $\tilde{\mathbf{D}} = \exp(\tilde{\mathbf{L}})$. We proved that it yields excellent theoretical properties, such as the monotone interpolation of the determinants, and the prevention of the swelling effect.

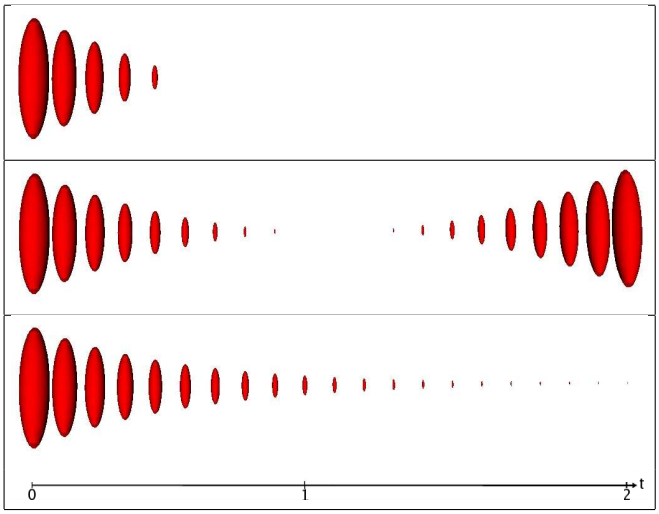

Fig. 1. **Geodesic shootings simulating a gradient descent. The x axis is the time t travelled along the geodesic ($\mathbf{t} \in [\mathbf{0}, \mathbf{2}]$). Top:** The Euclidean case. 2/3 of tensors are not positive definite and thus are not displayed. **Middle:** The Cholesky case: the null matrix is reached (exact middle value) and values beyond are the mirrored versions of the previous ones. **Bottom:** The Log-Euclidean case: all tensors are positive definite, and the null tensor is never reached.

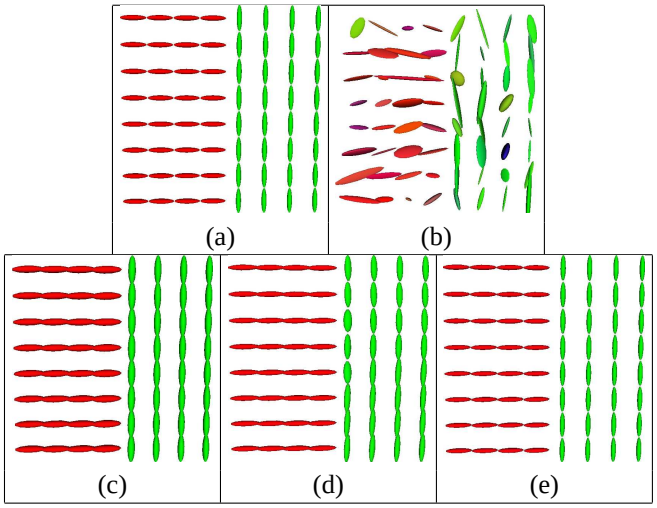

Fig. 2. **Anisotropic regularization of a noisy tensor field. (a):** Original field. **(b):** Noisy field. **(c):** Euclidean regularization. **(d):** Cholesky regularization. **(e):** Log-Euclidean regularization. Note the swelling effect in the Euclidean and Cholesky cases.

Advantages of the Log-Euclidean Framework:

To compare the use of a LE metric with the standard Euclidean framework and the more elaborate Cholesky decomposition, we believe that understanding the structure of the space is important. Let us illustrate this by realizing a geodesic shooting in the tensor space endowed with an Euclidean, Cholesky and Log-Euclidean structure. It consists in following the geodesic starting at one given tensor in a given direction during a certain time. In the Euclidean case, one simply computes: $\mathbf{D_e}(t) = \mathbf{D} + t\dot{\mathbf{D}}_\mathbf{e}$. In Figure 1, we chose a diagonal tensor $\mathbf{D} = ((4,0,0),(0,1,0),(0,0,1))$, a diagonal tangent vector $\dot{\mathbf{D}}_\mathbf{e} = ((-8,0,0),(0,-2,0),(0,0,-2))$ and $t \in [0,2]$. A "geodesic" with the Cholesky factors is given by a straight line in the space of lower triangular matrices: $\mathbf{L(t)} = \mathbf{L} + t\dot{\mathbf{L}}$. The resulting tensor curve is simply $\mathbf{D_c}(t) = \mathbf{L(t)L(t)}^\top$. To start at the same point with the same tangent vector as in the Euclidean case, $\mathbf{L}$ is the Cholesky factor of $\mathbf{D}$ and $\dot{\mathbf{L}}$ is solution of $\dot{\mathbf{L}}\,\mathbf{L}^\top + \mathbf{L}\,\dot{\mathbf{L}}^\top = \dot{\mathbf{D}}_\mathbf{e}$. In our case, we obtain $\dot{\mathbf{L}} = ((-2,0,0),(0,-1,0),(0,0,-1))$. With the LE metric, the geodesic is finally: $\mathbf{D_{log}} = \exp(\log(\mathbf{D}) + t\dot{\mathbf{D}}_\mathbf{log})$, where tangent vectors are given in the diagonal case by $\dot{\mathbf{D}}_\mathbf{log} = \exp(\mathbf{D})^{-1}\dot{\mathbf{D}}_\mathbf{e}$. Results of the geodesic shootings are displayed in Fig. 1.

In the Euclidean case, as expected, one quickly reaches the boundaries and non-displayed values are actually non-positive tensors. The Cholesky case is algorithmically well posed as non-positive matrices do not appear. However, we still reach zero eigenvalues, i.e. the null matrix appears, and values beyond it are the mirrored versions of the first ones. This means that the null matrix is reached on the trajectory during a gradient descent. Moreover, one may question the physical meaning of the mirrored values obtained beyond the null matrix. In the LE case, the null matrix is never reached and all tensors are by nature positive definite.

As a second advantage, the Log-Euclidean framework completely overcomes the *swelling effect* which can be observed in both Euclidean and Cholesky cases, and is illustrated in Fig. 2 with the example of the regularization. This effect causes tensors to grow after a processing. We generated a synthetic noisy tensor field and applied the anisotropic regularization of Sec. III-B with 3 different ways to process tensors. First, we used the Euclidean calculus, second we computed on the Cholesky factors and third we used a LE metric. Both Euclidean and Cholesky regularization suffer from the swelling effect, i.e. the denoised tensors are larger than the original values, the effect being less pronounced in the Cholesky case. With the LE framework, the swelling effect vanishes and the tensors are correctly denoised.

To quantify the benefit of the Log-Euclidean framework, we computed the root mean square error (RMSE) between the restored and original fields. Not to influence one particular metric, we computed the RMSE for the three metrics: Euclidean, Cholesky and LE. Results are summarized in Table I. Whatever the metric, the LE framework gives results quantitatively better than the 2 other frameworks.

To conclude, we believe that the Log-Euclidean framework is well adapted to the processing of diffusion tensors: it not only overcomes the limitations of the Euclidean calculus (negative eigenvalues), but it also removes the swelling effect which can be observed in both Euclidean and Cholesky frameworks. Consequently, we choose this family of metrics

TABLE I

RMSE between the restored and original tensor fields. FOR THE THREE GIVEN METRICS, THE LOG-EUCLIDEAN REGULARIZATION PRODUCES RESULTS THE CLOSEST TO THE ORIGINAL DATA.

	Euc. Reg.	Chol. Reg.	LE Reg.
Euclidean RMSE	0.228	0.172	0.051
Cholesky RMSE	0.152	0.092	0.015
Log-Euclidean RMSE	0.532	0.313	0.111

to solve our problem of joint estimation and smoothing of clinical DTI.

III. Joint Estimation and Regularization of Clinical DTI

The joint estimation and regularization of DTI can be tackled by a variational formulation, i.e. one has to minimize the energy functional:

$$E(\mathbf{L}) = \frac{1}{2}\text{Sim}(\mathbf{L}) + \frac{\lambda}{2}\text{Reg}(\mathbf{L}), \quad (1)$$

with Sim(.) being the data attachment term (estimation) and Reg(.) being the regularization term. In a statistical setting, Sim(.) is usually the log-likelihood of the measurements knowing the parameters, while Reg(.) is the prior knowledge on the parameters. λ is a normalization factor between the two terms. To fully make use of the advantages of the LE framework, we directly parameterize the diffusion tensor by its logarithm $\mathbf{L} = \log(\mathbf{D})$.

A. Three Least-Squares Criteria for DTI Estimation

The data attachment term can obviously take different forms. First, we propose a least-squares criterion on the logarithm of the DWI intensities, which is commonly used for tensor estimation. It corresponds to a maximum likelihood estimator when the noise is log-Gaussian (ML log-Gaussian). Second, we propose a least-squares criterion acting directly on the DWI intensities. In that case, the estimator corresponds to a ML with a Gaussian noise model (ML Gaussian). Third, we propose a new ML criterion handling the Rician nature of the noise (ML Rician).

1) Gaussian Noise on the Logarithm of the DWI Intensities: The linearized version of the Stejskal and Tanner diffusion equation system gives us the following energy to minimize, in the least-squares sense:

$$\text{Sim}_{log}(\mathbf{L}) = \int_\Omega \sum_{i=1}^{N} \left(\log\left(\frac{S_0}{\hat{S}_i}\right) - b\mathbf{g_i}^\top \exp(\mathbf{L})\mathbf{g_i} \right)^2, \quad (2)$$

where N is the number of encoding gradients. In order to minimize the criterion of Eq. [2], we need to differentiate it. The differentiation is easy to perform in the LE framework and gives:

$$\nabla\text{Sim}_{log}(\mathbf{L}) = -2\left(\log\left(\tfrac{S_0}{S_i}\right) - b\mathbf{g_i}^\top \exp(\mathbf{L})\mathbf{g_i}\right) \times \\ \times \partial_{\mathbf{G_i}} \exp(\mathbf{L}),$$

with $\mathbf{G_i} = \mathbf{g_i}\mathbf{g_i}^\top$ (see Appendix I for a practical implementation of the directional derivatives of the matrix exponential $\partial_{\mathbf{G_i}} \exp(\mathbf{L})$). Finally, the minimization is achieved through a simple first order gradient descent: $\mathbf{L_{t+1}} = \mathbf{L_t} - dt\nabla\text{Sim}_{log}(\mathbf{L_t})$. $\mathbf{L}$ belongs to a vector space and this evolution equation is actually a geodesic marching. After convergence, one simply needs to exponentiate the vector $\mathbf{L}$ to obtain a tensor: $\mathbf{D} = \exp(\mathbf{L})$.

2) Gaussian Noise on the Original DWI Intensities: Assuming a Gaussian noise on the original DWI intensities, the ML estimation boils down to a least-squares estimation of the tensor field from the images themselves rather than from their logarithm versions. For more clarity, we denote the predicted DWI intensity by $S_i(\mathbf{L}) = S_0 \exp(-b\mathbf{g_i} \exp(\mathbf{L})\mathbf{g_i}^\top)$. This gives the following criterion [7]:

$$\text{Sim}_{signal}(\mathbf{L}) = \sum_{i=1}^{N} \int_\Omega (S_i(\mathbf{L}) - \hat{S}_i)^2. \quad (3)$$

The differentiation of Eq. [3] gives:

$$\nabla\text{Sim}_{signal}(\mathbf{L}) = -2b \sum_{i=1}^{N} \left(S_i(\mathbf{L}) - \hat{S}_i\right) \times \\ \times S_i(\mathbf{L})\, \partial_{\mathbf{G_i}} \exp(\mathbf{L}). \quad (4)$$

Similarly to Eq. [2], the minimization is achieved through a first order gradient descent.

3) A Maximum Likelihood Estimator for a Rician Noise: In this section, we consider that the noise in MR images is Rician, i.e. the measured magnitude of the DWI can be modeled as:

$$\hat{\mathbf{S}}_\mathbf{i} = \sqrt{[\mathbf{S_i} + \mathbf{N_{re}}(\mathbf{0}, \sigma)]^\mathbf{2} + \mathbf{N_{im}}(\mathbf{0}, \sigma)^\mathbf{2}}$$

where $N_{re}(0,\sigma)$ and $N_{im}(0,\sigma)$ are independent Gaussian noises acting respectively on the real and imaginary part of the signal. The square magnitude of the observation $\hat{S}_i$ is:

$$\hat{S}_i^2 = (S_i + N_{re}(0,\sigma))^2 + N_{im}(0,\sigma)^2.$$

Taking the mean of the last expression gives:

$$\begin{aligned} E\left[\hat{S}_i^2\right] &= E\left[(S_i + N_{re}(0,\sigma))^2\right] + E\left[N_{im}(0,\sigma)^2\right] \\ &= E\left[S_i^2\right] + 2\sigma^2. \end{aligned}$$

Therefore, the observed square magnitude $\hat{S}_i^2$ is a non-central chi-square random variable. In this case, the DWI signal (not the squared version) is shifted by approximatively $\sigma^2/(2S_i)$ [8]. This means that the Rician noise induces a shrinking effect of tensors: The DWI signal tends to be greater than it should be, and the resulting tensors tend to be smaller than they actually are (a higher signal means a lower diffusion). This effect is even more obvious when the variance of the noise is high. To correct for this shrinking effect, we propose the ML estimator for the Rician noise.

For a Rician noise of variance σ^2 on the data, the pdf of the measured signal $\hat{S}$ knowing the expected signal S is [8]:

$$p(\hat{S}|S) = \frac{\hat{S}}{\sigma^2} \exp\left(-\frac{\hat{S}^2 + S^2}{2\sigma^2}\right) I_0\left(\frac{S\hat{S}}{\sigma^2}\right), \quad (5)$$

where I_0 is the modified 0^{th} order Bessel function of the first kind. The ML estimator for the pdf of Eq. [5] is:

$$\text{Sim}_{ML}(\mathbf{L}) = -\sum_{i=1}^{N} \log\left(p\left(\hat{S}_i|S_i(\mathbf{L})\right)\right). \quad (6)$$

The differentiation of Eq. [6] gives:

$$\nabla\text{Sim}_{ML}(\mathbf{L}) = -1/\sigma^2 \textstyle\sum_{i=1}^{N}(S_i(\mathbf{L}) - \alpha\hat{S}_i) \times \\ \times S_i(\mathbf{L})\, \partial_{\mathbf{G_i}} \exp(\mathbf{L}),$$

with $\alpha = I_0'/I_0(\hat{S}_i S_i/\sigma^2)$ (see Appendix II for a practical implementation of α). The formulation is similar to the gradient of Eq. [4], except that a correcting factor α depending on the signal and the noise variance appears. A simple estimator of the noise variance is based on the following. Typical MRIs include regions outside of the patient. Considering the fact that the square magnitude of such regions is null, taking its mean gives us an estimation of $2\sigma^2$.

B. An Anisotropic Regularization Term

Let us now turn to the prior on tensors: We expect the tensor field to vary spatially slowly within homogeneous regions (where the spatial gradient is small), while it can drastically change at the boundaries of these regions. The log-probability of such a prior can be efficiently represented by the ϕ-functionals usually used for anisotropic regularization: $\text{Reg}(\mathbf{L}) = \int_\Omega \phi\left(\|\nabla \mathbf{L}\|\right)$. The ϕ-function gives an anisotropic behavior to the regularization, i.e., it will preserve the edges of the tensor field while smoothing homogeneous regions. Note that this type of regularization using LE metrics was already proposed in [14], but without being coupled to an estimation criterion. Moreover, we give here the implementation details of such a regularization procedure using a finite differences strategy.

As we are working on a vector space, the gradient of the regularization criterion can be expressed as follows, using $\psi(s) = \phi\prime(s)/s$:

$$\begin{aligned} \nabla \text{Reg}(\mathbf{L}) &= -2\text{div}\left(\psi\left(\|\nabla \mathbf{L}\|\right)\nabla \mathbf{L}\right) \\ &= -2\psi\left(\|\nabla \mathbf{L}\|\right)\Delta \mathbf{L} - 2\sum_{i=1}^{3}\partial_i\left(\psi\left(\|\nabla \mathbf{L}\|\right)\right)\partial_i \mathbf{L}. \end{aligned} \quad (7)$$

For the experiments, we used $\phi(s) = 2(1+s^2/\kappa^2)^{1/2} - 2$ and $\psi(s) = (1+s^2/\kappa^2)^{-1/2}$ as in [12]. κ can be seen as a normalization factor for the gradient. The key for the numerical implementation is the computation of the matrix and scalar fields $\partial_i \mathbf{L}$, $\Delta \mathbf{L}$ and $\|\nabla \mathbf{L}\|$ of $\mathbb{R}^3$. Using a finite difference scheme, these are simply:

$$\begin{aligned} \partial_i \mathbf{L}(x) &= \frac{\mathbf{L}(\mathbf{x}+\mathbf{x_i}) - \mathbf{L}(\mathbf{x}-\mathbf{x_i})}{2\|x_i\|}, \\ \Delta \mathbf{L}(\mathbf{x}) &= \sum_{i=1}^{3}\frac{\mathbf{L}(\mathbf{x}+\mathbf{x_i}) - 2\mathbf{L}(\mathbf{x}) + \mathbf{L}(\mathbf{x}-\mathbf{x_i})}{\|x_i\|^2}, \\ \|\nabla \mathbf{L}(\mathbf{x})\|^2 &= \sum_{i=1}^{3}\|\partial_i \mathbf{L}(\mathbf{x})\|_{LE}^2, \end{aligned}$$

$\|.\|_{LE}$ being the Log-Euclidean metric of [14]. Finally, by combining the gradient of one of the criteria (Eq. [2], [3] or [6]) with the gradient of Eq. [7], one obtains the evolution equation of the joint estimation and smoothing of DTI:

$$\begin{aligned} \mathbf{L_{t+1}} &= \mathbf{L_t} - dt\nabla E(\mathbf{L_t}) \\ &= \mathbf{L_t} - dt/2\left(\nabla \text{Sim}(\mathbf{L_t}) + \lambda \nabla \text{Reg}(\mathbf{L_t})\right). \end{aligned}$$

Of course, one has to take the exponential of the solution to obtain a tensor.

We call the full criterion a Maximum A Posteriori estimator, because a spatial prior (the regularization term) is taken into account. As a conclusion of this section, we derive three new potential estimators: the MAP log-Gaussian (ML log-Gaussian estimator + regularization), the MAP Gaussian (ML Gaussian + regularization) and finally the MAP Rician (ML Rician + regularization). We now investigate the effects of these estimators on synthetic and real datasets.

IV. Results on Synthetic and Clinical Data

To illustrate the benefits of our methodology, we first perform 7 types of estimation on synthetic data: a classic estimation with an algebraic resolution (Classic), the ML log-Gaussian (Eq. [2]), the ML Gaussian (Eq. [3]) and the ML Rician (Eq. [6]). Then, the regularization term is added to turn each ML estimator into MAP estimations (MAP log-Gaussian, MAP Gaussian, MAP Rician), which gives a total of 7 different estimations. Second, we apply the same methodology on 2 clinical datasets, of medium and low quality. Results are presented and discussed in the sequel.

A. Synthetic Data

We generated a synthetic $16\times16\times16$ tensor field containing two homogeneous regions with anisotropic tensors (Fig. 3 left) as in [9]: the first region (R1) contains tensors whose coefficients are: $(0.970, 0.0, 1.751, 0.0, 0.0, 0.842)$ stored this way: $(d_{xx}, d_{xy}, d_{yy}, d_{xz}, d_{yz}, d_{zz})$. The second region (R2) contains tensors defined as: $(1.556, 0.338, 1.165, 0.0, 0.0, 0.842)$. A corresponding S_0 image is created with a constant value of 10. The DWI are artificially produced using the Stejskal & Tanner equation with 6 diffusion gradients simulating the real data of Sec. IV-B. Finally, a Rician noise is added to each simulated DWI, including the S_0, with three standard deviations: 0.5 (SNR$\simeq$ $10dB$), 1.0 (SNR$\simeq$ $8dB$) and 1.5 (SNR$\simeq 6dB$). Fig. 3 b, c and d show a slice of a DWI for the three levels of noise. Parameter κ for smoothing was set to 0.05. We found that this value gives good smoothing results while keeping most of the transitions. Going up to 0.1$\sim$0.5 will smooth a little more the data, but won't preserve much the interfaces in the tensor field, while going down to 0.01 will preserve almost everything and thus result in less smoothing. Moreover, as LE metric are similitude-invariant, κ does not depend on the scale of the tensors, and the given values can be used with any type of dataset. We set $\lambda = 1.0$ and $dt = 100$ iterations, which gives correct results in terms of speed of convergence, stability of the gradient descent, and influence of the regularization for the MAP estimators. A good range for λ is $[0.25, 1]$. Results of the Classic, ML log-Gaussian, ML Gaussian and ML Rician estimations are shown respectively in Fig. 4 b, c, d and e. Results of the MAP log-Gaussian, MAP Gaussian and MAP Rician estimations are presented respectively in Fig. 4 f, g, h.

To quantitatively compare the methods, we computed the mean error, the variance, the minimum and maximum errors between each estimation and the original data with the LE metric. Results are summarized in Table II. Since non positive tensors appear with a classical estimation, the LE metric

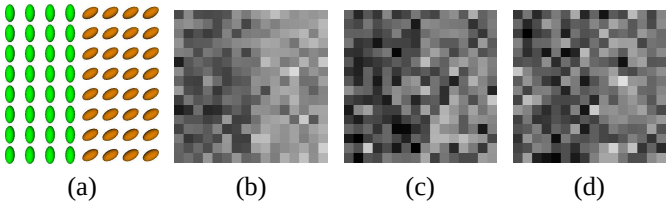

Fig. 3. **A slice of a generated DWI with $\sigma_n = 0.5$ (b), $\sigma_n = 1.0$ (c), $\sigma_n = 1.5$ (d).** The encoding gradient is $\mathbf{g} = (-1, 1, 0)^\top$.

gives an infinite error. With the ML log-Gaussian estimation, all tensors are by nature positive definite. However, negative tensors in a classical estimation are turned into tensors whose eigenvalues are as close as possible to zero. Indeed, the energy is minimum when these tensor eigenvalues are negative. This results in a slow convergence and high errors. On the contrary, the ML Gaussian estimator prevents the tensors to degenerate. Finally, the ML Rician estimator allows us to correct for the bias induced by the noise in each DWI and the shrinking effect vanishes, resulting in better quantitative results. The MAP estimators present the nice feature to preserve discontinuities between tensors, even when the tensors shapes are similar. This is an interesting feature when one would like, for instance, to smooth brain DTI: regions delimiting different fiber tracts have tensors about the same size but with different orientations. With our MAP estimators, these limits will be preserved.

Effects of these estimators on FA and ADC are detailed in Tab. II (last two columns). One notices that FA increases with noise, as reported in [18], [19]. ML Gaussian and Rician estimators are slightly less sensitive to noise w.r.t. FA than a Classic or ML Log-Gaussian estimator (the growth in FA is lower). The main advantage comes with the regularization (MAP estimators) which lower the FA compared to ML estimators. Conversely, ADC is underestimated when noise increases, which is a consequence of the shrinking effect. The ML and MAP Rician estimators suffer less from this shrinking, managing to recover almost the loss in tensor size.

To even more illustrate the shrinking effect, we evaluated the percentage of tensor volume lost during the estimation

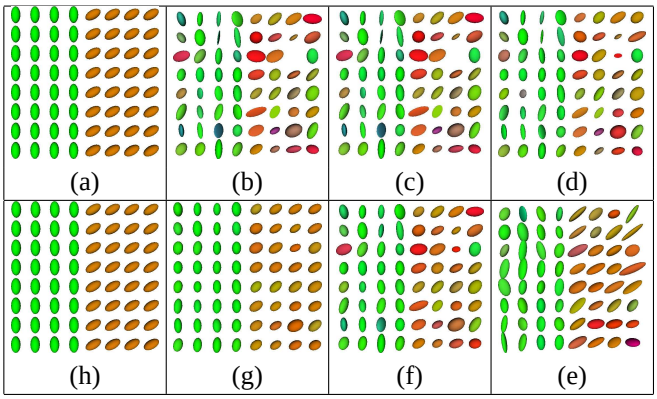

Fig. 4. **2D slices of 7 different estimations of a noisy synthetic DTI dataset ($\sigma_n = 1.0$). (a):** The 3D synthetic field. **(b):** The classic estimation (non-displayed ellipsoids correspond to non-positive tensors). **(c):** The ML log-Gaussian, **(d):** The ML Gaussian, **(e):** The ML Rician, **(f):** The MAP log-Gaussian, **(g):** The MAP Gaussian, **(h):** The MAP Rician.

(the tensor volume is its determinant) for the three levels of noise. Results are presented in Table III. The shrinking effect is obvious when the SNR is low. As depicted in Table III, the ML Rician estimator corrects for this effect, even when the noise variance is high.

These experiments show that using the ML estimator for the correct noise model helps to correct for the shrinking effect one can observe with a Rician noise and a low SNR. Moreover, the anisotropic regularization enforces the spatial correlation while preserving discontinuities of the diffusion tensor field, making our estimators suitable for clinical datasets with low SNRs.

B. Clinical Data

We tested the methods on 2 clinical datasets of medium and low quality. First, we used a brain dataset (Fig. 5) acquired on a 1.5T scanner with 7 encoding gradients (Basser sequence [1], b-value of 1000 s.mm^{-2}). The image dimensions are $128 \times 128 \times 30$ and the spatial resolution is $1.875 \times 1.875 \times 4\,\text{mm}^3$. Second, we used an experimental acquisition of the spinal cord on a 1.5T scanner[1] (Fig. 6 and 10) obtained with the same 7 encoding gradients and b-value as previously. The dimensions are $128 \times 128 \times 24$ (acquisition is coronal) with a spatial resolution of $1.4 \times 1.4 \times 1.4\,\text{mm}^3$. This new type of acquisition is currently actively investigated in clinical research (e.g., see [20]) and is difficult to perform. Indeed, the patient often cannot stay long enough in the scanner due to pathology. Moreover, the small entrance of the scanner forces the patient to have an uncomfortable position, and the scanning time must be shortened. Finally, the coil cannot be perfectly adapted to the body as it is for the head. The images are consequently much noisier than for the brain MRI. Note that these datasets were actually collected in a clinical environment: the brain dataset corresponds to a patient with a tumor, and the spinal cord one was acquired to check for possible compressions [20]. Estimation of the SNR in the brain dataset gives $\sigma = 12$ (SNR$\simeq 8dB$). Estimation of the SNR of the spinal cord dataset gives $\sigma = 14$ (SNR$\simeq 6dB$). Parameters used for the estimation are: $\kappa = 0.05$, $\lambda = 0.25$, $dt = 1.0$ and 100 iterations. Each of the ML and MAP log-Gaussian, Gaussian and Rician estimations took about 12 minutes to run on a PC with a Pentium M at 2GHz with 1 Go of memory.

Figures 5 show a closeup of the splenium region and nearby. We clearly see that the missing tensors in a Classic estimation (fig. 5 c) that are not positive-definite are not completely replaced with an ML log-Gaussian estimation (fig. 5 d) due to degenerate tensors. Using a ML Gaussian estimation (fig. 5 e) and the ML Rician (fig. 5 f) results in a field where all tensors are positive-definite. Below, we investigate quantitatively the effects of these estimators on the tensor field.

To evaluate the quality of the tensor fields estimated, we computed the mean apparent diffusion coefficient (ADC), the mean fractional anisotropy (FA) and the mean volume (VOL) in 2 distinct regions with different diffusion properties. First, we chose the ventricles, where the diffusion is high but

[1]The authors would like to thank Denis Ducreux, MD, for providing the brain and spinal cord DTI dataset.

isotropic, and second, we chose the corpus callosum (including the splenium and genu), where the diffusion is known to be restricted by neural fibers, thus exhibiting high FA values. Regions were segmented manually on axial slices. Results for the 7 estimations are presented in Table IV. As expected, the ML/Map Rician estimations produce the greatest volumes (we found an increase of about 14%), and the highest ADC values. A good property is that in anisotropic regions like the corpus callosum, the gain of volume does not induce a loss of anisotropy (see Table IV right, last column).

Finally, adding the regularization term (fig. 5 g, h and i) not only smooths isotropic regions like the ventricles without blurring the nearby splenium tract, but also replaces non-positive tensors of the ML log-Gaussian estimation by positive ones. Other experiments showed that the same comments apply for the interface between grey and white matter. FA maps are shown in Fig. 7 first row. However, the effect on the tensor volume, ADC and FA is more difficult to analyze. In isotropic regions like the ventricles, the nearby anisotropic regions have a small influence during the regularization process. Thus, tensors on the boundaries are corrupted by small anisotropic ones. Consequently, mean tensor volumes in the

TABLE II

Estimation of a synthetic dataset with 7 methods. RMSE ARE CALCULATED WITH THE LOG-EUCLIDEAN METRIC. FA AND ADC COLUMNS INDICATE MEAN VALUES OF THE ESTIMATIONS. THE PERCENTAGE IS THE RELATIVE DIFFERENCE TO THE TRUE MEAN VALUES (FA$= 0.39$, ADC$= 3.56\,10^{-3} mm^2 s^{-1}$). FA INCREASES WITH NOISE. USING MAP ESTIMATORS PERMITS TO CANCEL THIS EFFECT. CONVERSELY, ADC IS UNDERESTIMATED WHEN NOT USING THE ML/MAP RICIAN ESTIMATORS.

$\sigma_n = 0.5$

	NPT	Mean Error	Variance	Min Error	Max Error	FA	ADC ($10^{-3} mm^2 s^{-1}$)
Classic	7	∞	∞	0.077	∞	0.403 (2.7%)	3.567 (0.1%)
ML log-Gaussian	0	0.459	0.342	0.077	13.670	0.403 (2.7%)	3.567 (0.1%)
ML Gaussian	0	0.439	0.052	0.077	3.389	0.397 (1.2%)	3.534 (-0.8%)
ML Rician	0	0.250	0.007	0.051	0.666	0.402 (2.4%)	3.575 (0.3%)
MAP log-Gaussian	0	0.163	0.005	0.035	0.937	0.393 (0.1%)	3.573 (0.3%)
MAP Gaussian	0	0.141	0.003	0.032	0.374	0.385 (-1.9%)	3.517 (-1.3%)
MAP Rician	0	0.075	0.002	0.011	0.262	0.390 (-0.6%)	3.558 (-0.1%)

$\sigma_n = 1.0$

	NPT	Mean Error	Variance	Min Error	Max Error	FA	ADC($10^{-3} mm^2 s^{-1}$)
Classic	224	∞	∞	0.128	∞	0.428 (9%)	3.552 (-0.3%)
ML log-Gaussian	0	1.641	8.777	0.128	14.255	0.428 (9%)	3.552 (-0.3%)
ML Gaussian	0	1.086	0.718	0.128	5.413	0.410 (4.5%)	3.427 (-3.8%)
ML Rician	0	0.718	0.064	0.095	2.397	0.412 (5%)	3.58 (0.4%)
MAP log-Gaussian	0	0.584	0.051	0.011	2.032	0.409 (4%)	3.564 (0.1%)
MAP Gaussian	0	0.543	0.034	0.037	1.648	0.369 (-6%)	3.371 (-5.4%)
MAP Rician	0	0.120	0.004	0.017	0.545	0.390 (-0.6%)	3.534 (-0.8%)

$\sigma_n = 1.5$

	NPT	Mean Error	Variance	Min Error	Max Error	FA	ADC($10^{-3} mm^2 s^{-1}$)
Classic	717	∞	∞	0.205	∞	0.446 (14%)	3.423 (-4%)
ML log-Gaussian	0	3.518	22.921	0.205	14.422	0.446 (14%)	3.423 (-4%)
ML Gaussian	0	1.889	2.117	0.205	7.212	0.422 (7.5%)	3.207 (-10%)
ML Rician	0	1.525	0.542	0.261	6.332	0.441 (12%)	3.61 (1.3%)
MAP log-Gaussian	0	0.989	0.140	0.188	3.039	0.425 (8%)	3.434 (-3.6%)
MAP Gaussian	0	0.984	0.134	0.208	2.932	0.366 (-6.7%)	3.107 (-13%)
MAP Rician	0	0.394	0.046	0.033	1.889	0.410 (4.4%)	3.474 (-2.5%)

TABLE III

Illustration of the shrinking effect. MEAN VOLUMES ARE MEAN TENSOR DETERMINANTS OF EACH ESTIMATION. THE PERCENTAGE OF VOLUME LOSS INCREASES WITH THE NOISE VARIANCE. NOTE THAT THE ML RICIAN ESTIMATOR IS CORRECTING FOR THIS EFFECT.

$\sigma_n = 0.5$

	Original Data	ML log-Gaussian	ML Gaussian	ML Rician
Mean volume	1.43	1.41	1.39	1.43
Volume loss	NA	1.4%	2.9%	0.0%

$\sigma_n = 1.0$

	Original Data	ML log-Gaussian	ML Gaussian	ML Rician
Mean volume	1.43	1.25	1.15	1.42
Volume loss	NA	12%	19.2%	0.7%

$\sigma_n = 1.5$

	Original Data	ML log-Gaussian	ML Gaussian	ML Rician
Mean volume	1.43	1.11	0.81	1.40
Volume loss	NA	22%	43%	2%

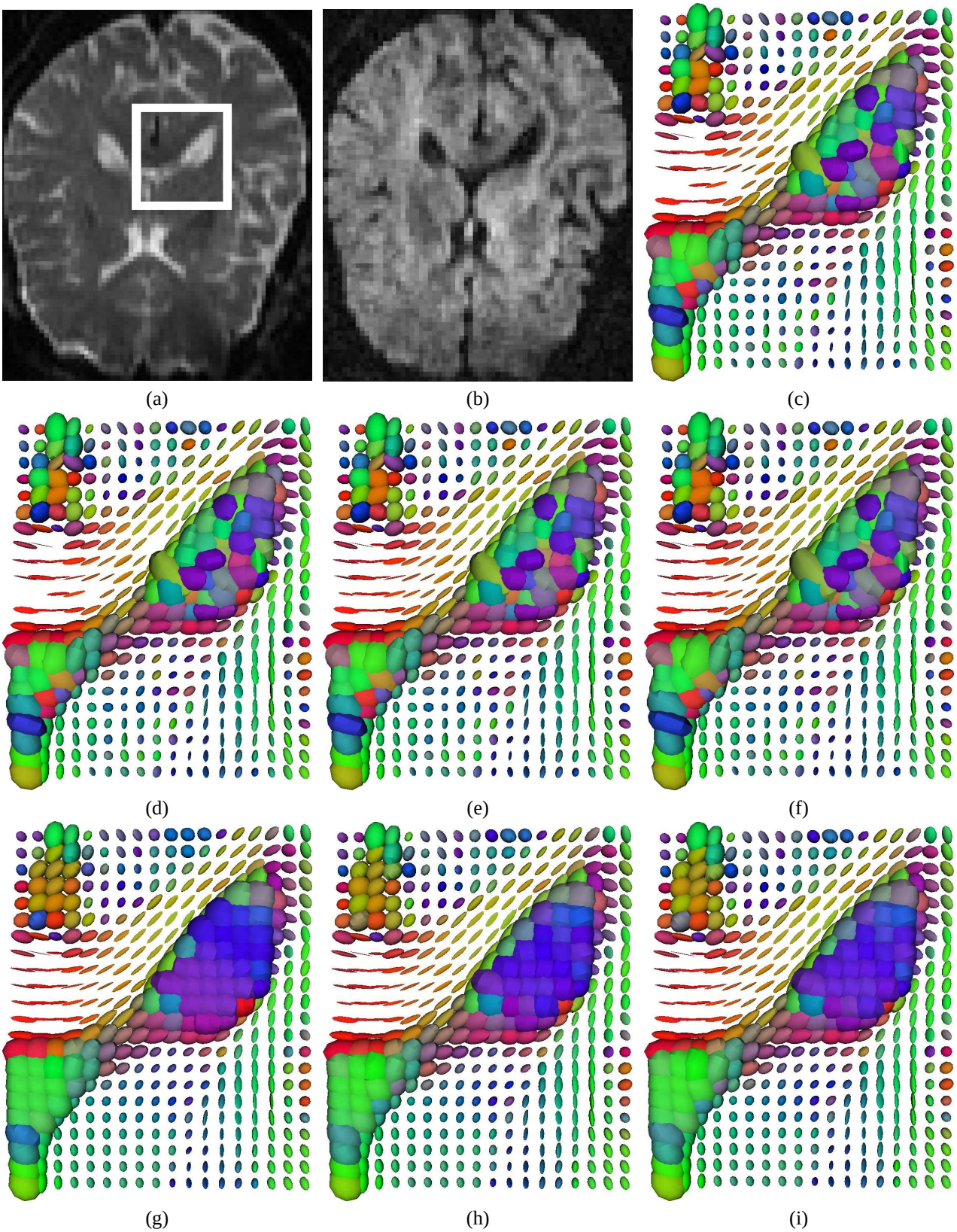

Fig. 5. **Tensor field estimation of a brain DTI dataset. (a):** The S_0 image. The other figures are a closeup on the region delimited by the white square. **(b):** The DWI corresponding to the encoding gradient $\mathbf{g} = (1, 0, 1)$. **(c):** Classic estimation. The color codes for the principal direction of tensors: **red:** left-right, **green:** anterior-posterior, **blue:** inferior-superior. Missing tensors in the splenium region are non-positive. **(d):** ML log-Gaussian. Some tensors are still missing because their eigenvalues are very close to zero. **(e):** ML Gaussian. All tensors remain positive definite. **(f):** ML Rician. All tensors are positive definite and are slightly bigger. **(g):** MAP log-Gaussian. The regularization term prevents the appearance of non-positive tensors. Note that the boundary between the ventricles and the splenium was preserved. **(h):** MAP Gaussian and **(i):** MAP Rician. The three MAP estimators give very close results. Note that our color scheme imposes to choose an orientation for coloring a glyph, even in case of isotropic tensors. Thus, inside the ventricles where tensors are isotropic after regularization (Fig. g, h and i), the same blue color is chosen arbitrarly.

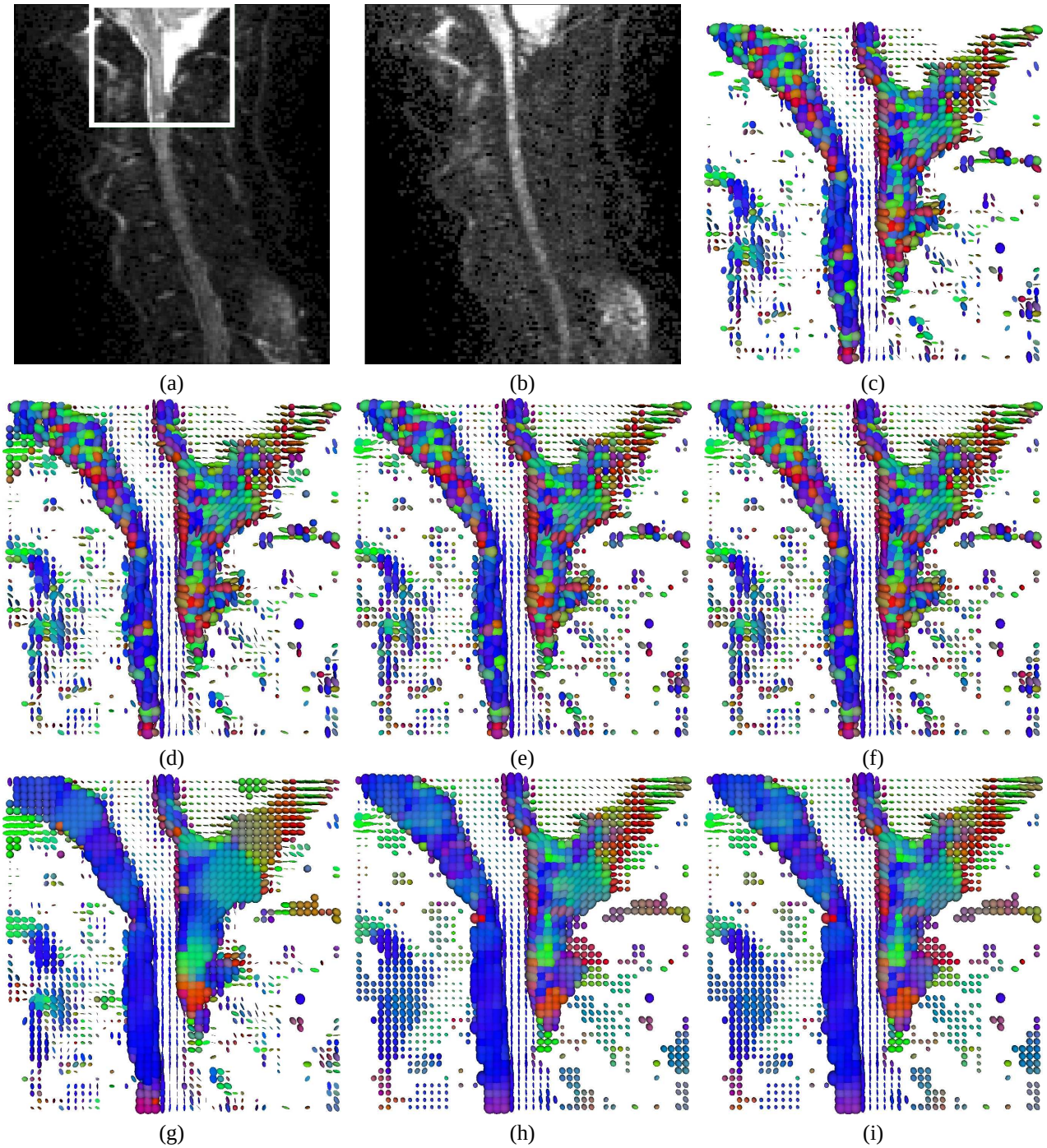

Fig. 6. **Tensor field estimation of a spinal cord dataset. (a):** The S_0 image. The other figures are a closeup of the region delimited by the white square. **(b):** The DWI corresponding to the encoding gradient $\mathbf{g} = (1, 0, 1)$. **(c):** Classic estimation. Many tensors are missing in and around the spinal cord. **(d):** ML log-Gaussian. Same tensors than with the Classic estimation are missing. **(e):** ML Gaussian. All tensors are positive definite and regions outside the spinal cord are coherent and show no artificial anisotropy. **(f):** ML Rician. **(g):** MAP log-Gaussian, **(h):** MAP Gaussian and **(i):** MAP Rician. The spinal tract is smoothed and boundaries with nearby isotropic regions are preserved.

ventricles with MAP estimators are lower than those without regularization. The same remark applies to the ADC (see Table IV left, second column). Conversely, in regions with anisotropic tensors, large isotropic neighbors may influence the results, leading to higher volume and ADC (Tab. IV right, first and second columns). Effects on FA maps are shown in Fig. 7 last row.

Figure 6 shows the results of the estimations of the spinal cord dataset. A closeup is made on the top of the spinal cord. The same remarks as for the brain dataset apply: the Classic estimation (Fig. 6 c) and the ML log-Gaussian (Fig. 6 d) lead to approximately the same results. Working with ML Gaussian and Rician estimators (Fig. 6 e and f) ensures that all tensors remain positive, with the advantage that the ML Rician estimator corrects for the shrinking effect: tensor volumes inside the spinal cord have grown by about 30%

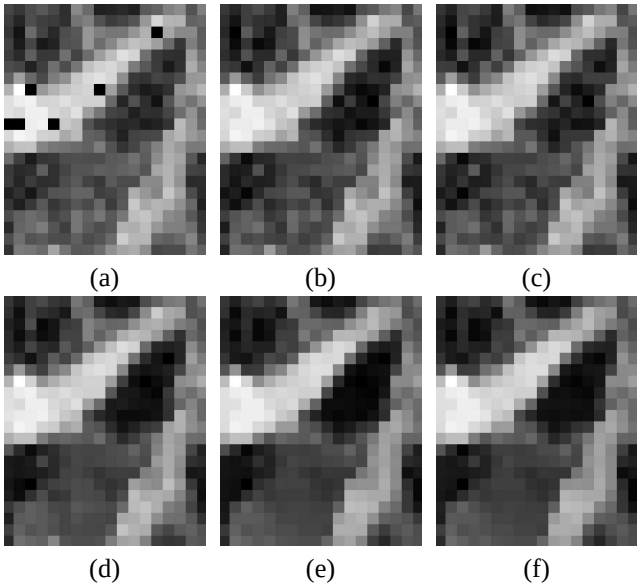

Fig. 7. **Fractional Anisotropy (FA) of the tensor fields obtained in Fig. 5. (a):** ML log-Gaussian estimation. Outliers (black dots) in the splenium are caused by degenerate tensors. **(b):** ML Gaussian. One notices no outliers. **(c):** ML Rician. Same as previously. **(d):** Map log-Gaussian, **(e):** Map Gaussian and **(f):** Map Rician. With the three MAP estimators, the FA contrast between the ventricles and the splenium is well enhanced.

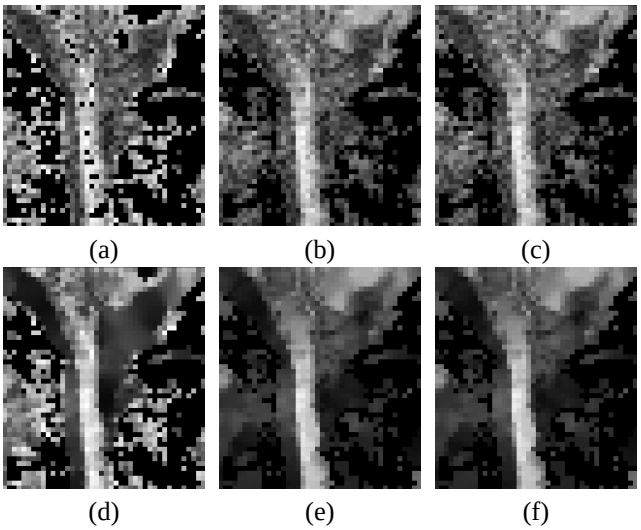

Fig. 8. **Fractional Anisotropy (FA) of the tensor fields obtained in Fig. 6. (a):** ML log-Gaussian estimation. Outliers (black dots) in the spinal cord are caused by degenerate tensors. **(b):** ML Gaussian. One notices no outliers. **(c):** ML Rician. Same as previously. **(d):** Map log-Gaussian, **(e):** Map Gaussian and **(f):** Map Rician. With the three MAP estimators, the FA contrast between the spinal cord and around is well enhanced.

compared to the Classic estimation (the spinal cord was manually segmented from the baseline image). We found a greater growth of tensor volumes in the spinal cord than in the brain (within the brain, tensors have grown on average by 10%). This difference can be explained by the noise level: the SNR of the spinal cord dataset is the lowest. Effects on FA maps are presented in Fig. 8 top row. The regularization term (Fig. 6 g, h and i) smooths the field while preserving the boundaries with the spinal tract.

With the MAP log-Gaussian estimation, some artificial anisotropic tensors appear: The FA map (Fig. 8 d) presents high values outside the spinal cord, and is thus noisier than those obtained with MAP Gaussian and Rician estimators (Fig. 8 e and f).

This exemplifies the importance of the choice of the noise model: For low quality data, considering a log-Gaussian noise may not be the right choice, even with a MAP estimation. Switching to Gaussian or Rician noise models can largely improve the quality of the tensor estimation.

We now study how the MAP Rician estimator impacts the quality of fiber tracking on these two datasets.

C. Improvement of Tractography

Tractography, or fiber tracking, is a process which runs at the end of the DTI processing pipeline. Among the numerous available methods for tracking fibers, we chose a relatively fast and easy to implement one [21] and show how the tracking can be improved by our variational estimation combined with regularization. Criteria for stopping the tracking are: a threshold on FA (if FA is too low, the tracking is stopped) and on the curvature (to forbid unlikely fibers having a high curvature). Prior to the tracking, tensor fields are resampled to obtain isotropic voxels: in general, the out-plane resolution is very low (e.g., the brain dataset here) and interpolating the tensors improves the regularity of the fibers. Resampling is interpreted as a weighted mean with trilinear coefficients.

TABLE IV

Quantitative comparison of 7 diffusion tensor estimations of a brain DTI dataset. FOR EACH ESTIMATION, THE MEAN VOLUME, ADC AND FA WERE EVALUATED IN THE VENTRICLES, AND THE CORPUS CALLOSUM. IN ISOTROPIC REGIONS LIKE THE VENTRICLES, ONE NOTICES THAT THE TENSOR VOLUME IS ON AVERAGE 14% (ML RICIAN) OR 10% (MAP RICIAN) LARGER THAN WITH A CLASSIC ESTIMATION. THE ADC SHOWS SLIGHTLY HIGHER VALUES WITH THE ML RICIAN ESTIMATION (4% GROWTH) AND MAP RICIAN (2% GROWTH) THAN WITH THE CLASSIC ESTIMATION. THE SAME REMARKS APPLY IN THE CORPUS CALLOSUM BUT IS LESS MARKED.

	Ventricles				**Corpus Callosum**		
	Volume	ADC (mm^2s^{-1})	FA		Volume	ADC (mm^2s^{-1})	FA
Classic	14.43	$6.63.10^{-3}$	0.26	Classic	0.63	$2.66.10^{-3}$	0.64
ML log-Gaussian	14.31	$6.62.10^{-3}$	0.25	ML log-Gaussian	0.63	$2.66.10^{-3}$	0.64
MAP log-Gaussian	13.82	$6.49.10^{-3}$	0.18	MAP log-Gaussian	0.61	$2.61.10^{-3}$	0.57
ML Gaussian	14.39	$6.63.10^{-3}$	0.25	ML Gaussian	0.63	$2.66.10^{-3}$	0.64
MAP Gaussian	14.18	$6.58.10^{-3}$	0.22	MAP Gaussian	0.65	$2.65.10^{-3}$	0.61
ML Rician	16.47	$6.87.10^{-3}$	0.27	ML Rician	0.65	$2.68.10^{-3}$	0.64
MAP Rician	15.94	$6.78.10^{-3}$	0.23	MAP Rician	0.67	$2.68.10^{-3}$	0.62

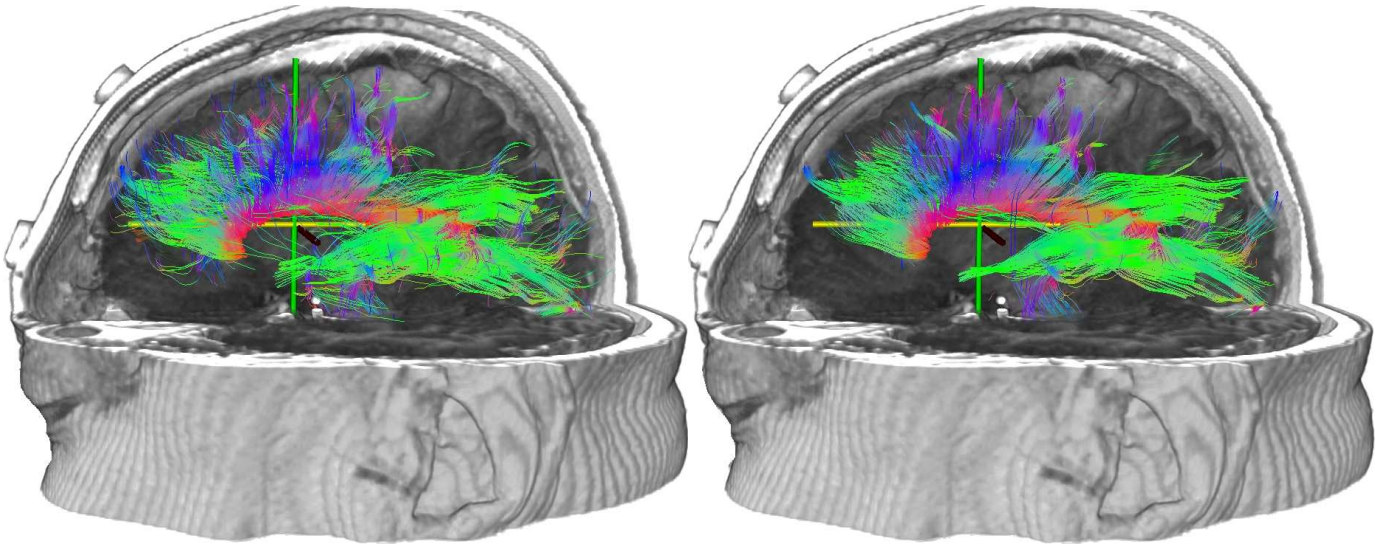

Fig. 9. **Improvement of fiber reconstruction.** A seed region was placed inside the corpus callosum. Results of the fiber reconstruction after a classical estimation (**Left**), and after the MAP Rician estimation (**Right**). Fibers are overlapped with a volume rendering of the T1 image.

Such a mean is computed in the logarithmic domain and then mapped back to the tensor space with the matrix exponential: $\mathbf{D} = \exp(\sum_{i=1}^{N} \omega_i \log(\mathbf{D_i}))$, where ω_i are classical trilinear weights. We showed in [14] that such an interpolation has good practical properties in the context of DT-MRI, compared to a Euclidean interpolation. We tracked the fibers from the tensor fields obtained after the Classic estimation plus resampling and the MAP Rician estimator plus resampling. The parameters used for the tracking are: FA threshold: 0.3, maximum angle of deviation: $90°$. Results of tracking in the brain and the spinal cord are shown in Fig. 9, and 10. With the MAP Rician estimator, the tracking is qualitatively much smoother in both cases and shows less dispersion. The smoothness of the tensor field leads to more regular and longer fibers: tracts that were stopped due to the noise are now fully reconstructed. The FA threshold used ensures that all fibers belong to white matter, and do not result from a tracking in CSF or grey matter.

V. Discussion and Conclusions

We presented a new methodology to process DTIs of medium and low quality (typical of clinical applications) through a joint estimation and regularization of the diffusion tensor field. In particular, the estimation, which assumes that the data are corrupted by a Rician noise, is achieved through

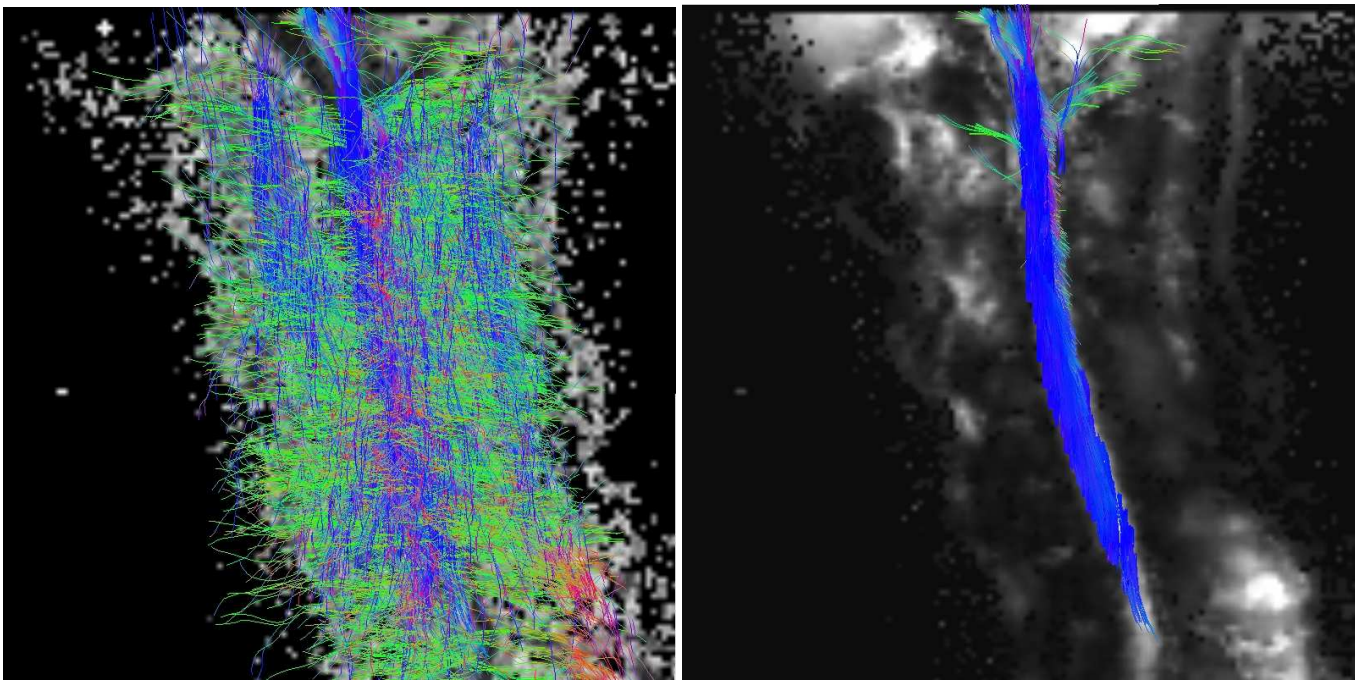

Fig. 10. **Spinal cord fiber tract reconstruction.** A region containing the spinal cord was used for the tracking. **Left:** The spinal cord reconstructed after the Classic estimation. **Right:** The same tract after our proposed variational framework. Fibers are overlapped with a slice of the FA map.

a maximum likelihood strategy adapted to the nature of this noise. This approach has the advantage to correct for the bias induced by the Rician noise in the DWI, and consequently not to underestimate the true volume of tensors (shrinking effect). Other estimation criteria which make the assumption of a Gaussian noise on the logarithm of the signal and on the signal itself are compared. These estimators are combined with an anisotropic regularization of the tensor field, so that transitions between homogeneous fiber tracts are preserved. To optimize these criteria, we use a Log-Euclidean metric that provides a fast and easy to use framework to process tensors. This tensor computing framework completely overcomes the limitations of the standard Euclidean calculus, and is well adapted to the processing of diffusion tensors. Results on synthetic data show that considering a ML estimation adapted to a Rician noise model corrects for the shrinking effect, while assuming other noise models results in a loss of tensor volume after estimation. Results on real clinical datasets show that the use of ML estimators can be valuable in clinical studies: The MAP estimations nicely smooths homogeneous regions without blurring transitions between different tract in the case of brain DTI. Moreover, with a dataset of low quality, we showed that the choice of the noise model is important, as outliers may persist even with the regularization. The ML Rician estimator turns out to be the best choice in that case. Finally, the promising improvement of the fiber reconstruction of these data shows that even clinical DTIs can be used for tractography.

Previously, diffusion tensor estimation was performed apart from tensor regularization. Recently, Basu et al. [10] regularized independently each DWI with a Rician noise removal approach, and estimated diffusion tensors afterwards. In this paper, we propose to do both together, i.e. to regularize the tensor field while estimating it from DWI with a Rician noise model. We believe that using such an approach allows to capture more information than considering each DWI individually. However, one should compare in practice the properties of DWI processing versus tensor processing, e.g. does DWI processing preserve FA, trace, and tensor volumes?

In the future, the questions of validation and reproducibility have to be answered. We could think of repeating scans of the same patient in various orientations in the scanner, and in various scanners, for this purpose. One also could think of using phantoms and histological data as in [22].

Finally, the observed qualitative impact on the tracking could be quantified, using for instance a dispersion measure of the fibers.

Appendix I
Practical Implementation of the Matrix Exponential Directional Derivative

In Sec. III, the directional derivative of the exponential $\partial_{\mathbf{G}} \exp(\mathbf{L})$ is used. For general matrices, one has to compute the series [14]:

$$\partial_{\mathbf{G}} \exp(\mathbf{L}) = \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \mathbf{L}^i \mathbf{G} \mathbf{L}^{k-i-1}. \tag{8}$$

In general, Eq. [8] cannot be simplified. However, in the case of symmetric matrices, the differential can take a much simpler form. We know that $\mathbf{G}$ and $\mathbf{L}$ are both symmetric matrices: In our application, $\mathbf{G}$ is the tensor product of a vector $\mathbf{g}$ with itself: $\mathbf{G} = \mathbf{g}\mathbf{g}^\top$, which is by construction symmetric, and $\mathbf{L}$ is the matrix logarithm of a diffusion tensor, therefore symmetric.

Let $\mathbf{L} = \mathbf{R}^\top \mathbf{S} \mathbf{R}$ be an eigen decomposition of $\mathbf{L}$. $\mathbf{S}$ is diagonal: $\mathbf{S} = \mathrm{diag}(s_1, s_2, s_3)$. From Eq. [8], we have:

$$\begin{aligned}
\partial_{\mathbf{G}} \exp(\mathbf{L}) &= \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \left(\mathbf{R}^\top \mathbf{S}^i \mathbf{R}\right) \mathbf{G} \left(\mathbf{R}^\top \mathbf{S}^{k-i-1} \mathbf{R}\right) \\
&= \mathbf{R}^\top \left(\sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \mathbf{S}^i \left(\mathbf{R}\mathbf{G}\mathbf{R}^\top\right) \mathbf{S}^{k-i-1} \right) \mathbf{R} \quad (9) \\
&= \mathbf{R}^\top \left(\partial_{\mathbf{R}\mathbf{G}\mathbf{R}^\top} \exp\left(\mathbf{S}\right)\right) \mathbf{R} \quad (10)
\end{aligned}$$

We have $\partial_{\mathbf{G}} \exp(\mathbf{L}) = \mathbf{R}^\top \partial_{\mathbf{R}\mathbf{G}\mathbf{R}^\top} \exp(\mathbf{S}) \mathbf{R}$. Let $\mathbf{M}$ be $\partial_{\mathbf{R}\mathbf{G}\mathbf{R}^\top} \exp\left(\mathbf{S}\right)$. We denote by $[\mathbf{M}]_{(l,m)}$ the coefficient (l,m) of matrix $\mathbf{M}$ (i.e. the matrix coefficient stored at row l and column m). $(l,m) \in \{1,2,3\}$ for diffusion tensors. From Eq. [9] we can express the coefficient (l,m) of $\mathbf{M}$ as:

$$[\mathbf{M}]_{(l,m)} = \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \left[\mathbf{S}^i \mathbf{R}\mathbf{G}\mathbf{R}^\top \mathbf{S}^{k-i-1}\right]_{(l,m)}.$$

As $\mathbf{S}$ is diagonal, we can further write:

$$\begin{aligned}
[\mathbf{M}]_{(l,m)} &= \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s_l^i \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} s_m^{k-i-1} \\
&= \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s_l^i s_m^{k-i-1}.
\end{aligned}$$

Now, we only need to express the series $\sum_{k=1}^{+\infty} 1/k! \sum_{i=0}^{k-1} s_l^i s_m^{k-i-1}$. If $s_l \neq s_m$, we have:

$$\begin{aligned}
\sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s_l^i s_m^{k-i-1} &= \sum_{k=1}^{+\infty} \frac{1}{k!} s_m^{k-1} \sum_{i=0}^{k-1} \left(\frac{s_l}{s_m}\right)^i \\
&= \sum_{k=1}^{+\infty} \frac{1}{k!} s_m^{k-1} \frac{1-(s_l/s_m)^k}{1-s_l/s_m} \\
&= \frac{1}{s_m - s_l} \sum_{k=1}^{+\infty} \frac{1}{k!} s_m^k \frac{s_m^k - s_l^k}{s_m^k} \\
&= \frac{\exp(s_m) - \exp(s_l)}{s_m - s_l}.
\end{aligned}$$

If $sl = s_m = s$, then the series can be even more simplified:

$$\begin{aligned}
\sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s^i s^{k-i-1} &= \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s^{k-1} = \sum_{k=1}^{+\infty} \frac{1}{k!} \left(k s^{k-1}\right) \\
&= \sum_{k=1}^{+\infty} \frac{s^{k-1}}{(k-1)!} = \sum_{k=0}^{+\infty} \frac{s^k}{k!} \\
&= \exp(s).
\end{aligned}$$

Finally, one can access the coefficient (l,m) of $\mathbf{M}$ as:

$$\begin{aligned}[M]_{(l,m)} &= \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} \frac{e^{s_m}-e^{s_l}}{s_m-s_l} \text{ if } s_l \neq s_m, \\ &= \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} e^{s_l} \text{ if } s_l = s_m. \end{aligned} \quad (11)$$

From Eq. [11], and Eq. [9], we deduce the directional derivative of the matrix exponential $\partial_{\mathbf{G}} \exp(\mathbf{L})$.

APPENDIX II
PRACTICAL IMPLEMENTATION OF $\alpha(x)$

In Sec. III-A.3, the coefficient $\alpha(x) = I_0'/I_0(x)$ is needed to be able to compute the gradients of the ML and MAP estimators. α takes its values between 0 and 1. A practical implementation of $\alpha(x)$ is given in the following.

The 0^{th} order modified Bessel function of the first kind $I_0(x)$ is expressed as follows: $I_0(x) = \sum_{k=0}^{+\infty} \frac{(x^2/4)^k}{k!^2}$. Consequently, $\alpha(x)$ can be written as:

$$\alpha(x) = \frac{I_0'(x)}{I_0(x)} = \frac{x}{2} \frac{\sum_{k=0}^{+\infty} \frac{\left(\frac{1}{4}x^2\right)^k}{k!(k-1)!}}{\sum_{k=0}^{+\infty} \frac{\left(\frac{1}{4}x^2\right)^k}{k!^2}}. \quad (12)$$

Unfortunately, Eq. [12] cannot be simplified. However, for small values of x (typically $x \leq 10$) the series converges very quickly, and in practice 10 terms of the series are sufficient for an accuracy of 1.10^{-12}, which is by far enough. For higher values of x, one may experience numerical stability issues. To prevent this, one may use the following Taylor expansion of $\alpha(x)$ valid for $x > 10$:

$$\begin{aligned}\alpha(x) \simeq\ & 1.0 - 0.5x^{-1} - 0.125x^{-2} - 0.125x^{-3} - 0.1953125x^{-4} \\ & -0.40625x^{-5} - 1.0478515625x^{-6} - 3.21875x^{-7} \\ & -11.46646118164x^{-8} - 46.478515625x^{-9} \\ & -211.276149749755x^{-10} - 1064.67822265625x^{-11}\end{aligned}$$

REFERENCES

[1] P. Basser, J. Mattiello, and D. L. Bihan, "MR diffusion tensor spectroscopy and imaging," *Biophysical Journal*, vol. 66, pp. 259–267, 1994.

[2] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache, "Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics," in *Proc. of ISBI'06*, Apr. 2006, pp. 786–789.

[3] P. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *Journal of Magnetic Resonance B*, vol. 103, no. 3, pp. 247–254, 1994.

[4] C.-F. Westin, S. Maier, H. Mamata, A. Nabavi, F. Jolesz, and R. Kikinis, "Processing and visualization for diffusion tensor MRI," *MedIA*, vol. 6, no. 2, pp. 93–108, 2002.

[5] R. Deriche, D. Tschumperlé, and C. Lenglet, "DT-MRI estimation, regularization and fiber tractography." in *Proc. of ISBI'04*, pp. 9–12.

[6] R. Salavador, A. Peña, D. Menon, T. Carpenter, J. Pickard, and E. Bullmore, "Formal characterization and extension of the linearized diffusion tensor model," *HBM*, vol. 24, pp. 144–155, 2005.

[7] L.-C. Chang, D. K. Jones, and C. Pierpaoli, "RESTORE: Robust estimation of tensors by outlier rejection," *MRM*, vol. 53, no. 5, pp. 1088–1095, 2005.

[8] J. Sijbers, A. den Dekker, P. Scheunders, and D. V. Dyck, "Maximum likelihood estimation of Rician distribution parameters," *IEEE TMI*, vol. 17, no. 3, June 1998.

[9] Z. Wang, B. Vemuri, Y. Chen, and T. Mareci, "A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI," *IEEE TMI*, vol. 23, no. 8, August 2004.

[10] S. Basu, T. Fletcher, and R. Whitaker, "Rician noise removal in diffusion tensor MRI," in *Proc. of MICCAI'06*, ser. LNCS 4190, M. Nielsen and J. Sporring, Eds. Springer, October 2006, pp. 117–125.

[11] O. Coulon, D. Alexander, and S. Arridge, "Diffusion tensor magnetic resonance image regularization," *MedIA*, vol. 8, no. 1, pp. 47–67, 2004.

[12] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, "Constrained flows of matrix-valued functions: Application to diffusion tensor regularization," in *Proc. of ECCV'02*, ser. LNCS 2350. Springer Verlag, 2002, pp. 251–265.

[13] X. Pennec, P. Fillard, and N. Ayache, "A Riemannian framework for tensor computing," *IJCV*, vol. 66, no. 1, pp. 41–66, January 2006.

[14] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-Euclidean metrics for fast and simple calculus on diffusion tensors," *MRM*, vol. 56, no. 2, pp. 411–421, August 2006.

[15] C. Lenglet, M. Rousson, and R. Deriche, "Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing," *JMIV*, vol. 25, no. 3, pp. 423–444, Oct. 2006.

[16] P. Fletcher and S. Joshi, "Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors." in *Proc. of CVAMIA and MMBIA*, ser. LNCS 3117, 2004, pp. 87–98.

[17] P. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly, "A rigorous framework for diffusion tensor calculus," *MRM*, vol. 53, pp. 221–225, 2005.

[18] A. W. Anderson, "Theoretical analysis of the effects of noise on diffusion tensor imaging," *MRM*, vol. 46, no. 6, pp. 1174–1188, December 2001.

[19] S. Skare, T. Li, B. Nordell, and I. M., "Noise considerations in the determination of diffusion tensor anisotropy," *MRI*, vol. 18, no. 6, pp. 659–669, July 2000.

[20] D. Facon, A. Ozanne, P. Fillard, J.-F. Lepeintre, C. Tournoux-Facon, and D. Ducreux, "MR diffusion tensor imaging and fiber tracking in spinal cord compression," *AJNR*, vol. 26, pp. 1587–1594, June-July 2005.

[21] P. Fillard and G. Gerig, "Analysis tool for diffusion tensor MRI," in *Proc. of MICCAI'03*, ser. LNCS 2878, no. 2, pp. 967–968.

[22] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Rivière, A. Constantinesco, D. L. Bihan, and J.-F. Mangin, "Fiber tracking in q-ball fields using regularized particle trajectories," in *Proc. of IPMI'05*, ser. LNCS 3565, G. E. Christensen and M. Sonka, Eds., vol. 1. Springer, July, pp. 52–63.