A Bayes Hilbert Space for Compartment Model Computing in Diffusion MRI

Aymeric Stamm, Olivier Commowick, Alessandra Menafoglio, Simon K Warfield

A Bayes Hilbert Space for Compartment Model Computing in Diffusion MRI

A. Stamm[1,2], O. Commowick[3], A. Menafoglio[4], and S.K. Warfield[2]

[1]CADS, Human Technopole, Milan, MI, Italy (`aymeric.stamm@htechnopole.it`)
[2]CRL, Boston Children's Hospital, Harvard Medical School, MA, USA
[3]Univ. Rennes 1, CNRS, Inria, Inserm, IRISA UMR 6074, VisAGeS U1228, France
[4]MOX, Dept. of Mathematics, Politecnico di Milano, Italy

Abstract. The single diffusion tensor model for mapping the brain white matter microstructure has long been criticized as providing sensitive yet non-specific clinical biomarkers for neurodegenerative diseases because (i) voxels in diffusion images actually contain more than one homogeneous tissue population and (ii) diffusion in a single homogeneous tissue can be non-Gaussian. Analytic models for compartmental diffusion signals have thus naturally emerged but there is surprisingly little for processing such images (estimation, smoothing, registration, atlasing, statistical analysis). We propose to embed these signals into a Bayes Hilbert space that we properly define and motivate. This provides a unified framework for compartment diffusion image computing. Experiments show that (i) interpolation in Bayes space features improved robustness to noise compared to the widely used log-Euclidean space for tensors and (ii) it is possible to trace complex key pathways such as the pyramidal tract using basic deterministic tractography thanks to the combined use of Bayes interpolation and multi-compartment diffusion models.

1 Introduction

Diffusion MRI is a unique in-vivo and non-invasive imaging technique that probes the microstructure of tissues by tracking water diffusion [5]. The diffusion process is a probability distribution that describes the random 3-dimensional displacements of water molecules due to thermal agitation. The signal recovered in diffusion MRI is the Fourier transform of the diffusion density evaluated at the magnetic field spatial gradient to which the imaged brain was subjected.

A practical approach to access the diffusion density pertains to resorting to parametric models that lead to an analytic expression of the observed diffusion signals. The first parametric model was introduced two decades ago in [2] and is called *single tensor* (ST) model. The ST model essentially assumes that the diffusion in the voxel is well characterized by a zero-mean Gaussian distribution with covariance matrix proportional to the so-called diffusion tensor. It has become widely used in clinical practice, which raised the demand for a solid mathematical framework for processing tensor images (estimation, smoothing, registration,

atlasing, statistical analysis). As a response, several frameworks have been proposed, including a Riemannian affine-invariant framework [9, 6] and the popular log-Euclidean space for tensors, embedding them in a Lie group structure [1].

However, because of the low spatial resolution in diffusion MRI, it has been shown that the ST model is an average of several diffusion processes arising from multiple populations of tissues in the voxel, ultimately leading to an inaccurate description of the microstructure in most parts of the brain white matter. In addition, even within a single homogeneous tissue population, non-Gaussian diffusion has been observed [13]. This has led to the extension of the ST model to mixture models, often called multi-compartment models (MCM), where the voxelwise diffusion signal is modeled as a linear combination of compartmental diffusion signals arising from underlying homogeneous diffusion processes [8]. To the best of our knowledge, the only mathematical framework available for multi-compartment image processing [12] is however limited to multi-tensor images (mixture of single-tensor signals, see [10]) as it relies on log-Euclidean geometry.

In this work, we present an alternative to the log-Euclidean space on tensors called the Bayes Hilbert space [3] for processing compartmental diffusion signals instead of tensors. It provides a unified framework that can accommodate analytic compartment diffusion signals of any type (Gaussian and non-Gaussian). Section 2 gives a general introduction to Bayes Hilbert spaces with motivation and setup in diffusion MRI. Section 3 describes a simulation study and a tractography application on real data to compare Euclidean, log-Euclidean and Bayes interpolation. Results commented in Section 4 show that Bayes interpolation features improved robustness to noise and yields better pyramidal tract reconstructions when combined with MCM-based deterministic tractography to account for streamline atlas priors when diffusion models cannot be locally trusted.

2   Theory

2.1   Bayes Hilbert Spaces and Compartmental Diffusion

The signal in diffusion MRI, hereafter diffusion signal, observed after the application of a diffusion gradient $\mathbf{q}$, where $\|\mathbf{q}\|^2 = b$ is the well known b-value, undergoes a signal decay starting from a baseline signal. The form of this decay depends on how water molecules diffuse in tissues in the vicinity of the spatial location where the signal is observed. In general, in a voxel composed of K different tissue structures, the compound diffusion signal reads:

$$S(\mathbf{q}) = \sum_{j=1}^{K} S_j(\mathbf{q}) = \sum_{j=1}^{K} S_{0j} A_j(\mathbf{q}), \quad S_{0j} > 0, \quad A_j \in [0, 1], \tag{1}$$

in which all the information pertaining to the microstructure (set of surrounding tissue structures) is grasped by the signal attenuations A_j's. This is the signal decomposition at the foundation of diffusion compartment imaging [8]. Hence, up to some multiplicative constant, the compartmental diffusion signal S_j can be

interpreted as the density of a probability measure ν_j on the space of diffusion gradients, which is absolutely continuous w.r.t. the Lebesgue measure λ such that $\frac{d\nu_j}{d\lambda}$ (Radon-Nykodym derivative w.r.t. the Lebesgue measure λ) is proportional to S_j. Furthermore, this multiplicative constant is related to the T_2 relaxation time of tissue j and thus not relevant for microstructure mapping.

To account for this, we propose to embed the probability measures induced by compartmental diffusion signals into the Bayes space $B^2(P)$ of (classes of equivalence of $\mathcal{B}(\mathbb{R}^3)$-finite) measures on the Borel space $(\Omega, \mathcal{B}(\Omega))$, with square-integrable log-densities w.r.t. a reference measure P [3], which reads:

$$B^2(P) = \left\{ \nu : \int_{\mathbb{R}^3} \log^2 \left(\frac{d\nu}{dP} \right) dP < \infty, \nu > 0 \right\}. \tag{2}$$

In this view, if a diffusion signal S_{j1} (or, equivalently a probability measure ν_1 induced by S_{j1}) carries a piece of information about the structure of tissue j, then a diffusion signal S_{j2} proportional to S_{j1} does not carry additional information about the structure of tissue j. Hence, S_{j1} and S_{j2} are regarded as equivalent for microstructure mapping. This key property, known as *scale invariance*, is accounted for in the Bayes space $B^2(P)$ by the induced classes of equivalence. Not every Bayes space $B^2(P)$ however can accommodate compartmental diffusion signals because, depending on the choice of the reference measure P, the logarithm of compartmental diffusion signals might not be square-integrable. The choice of the reference measure of the Bayes space is thus critical. Provided that an appropriate reference measure P exists, embedding diffusion signals in $B^2(P)$ provides a unified framework for processing any type of analytic microstructure compartment models [8] at the cost of numerical integrations.

The space $B^2(P)$ is a vector space when endowed with the perturbation and powering operations $(\oplus, \odot)$, defined for $\nu_1, \nu_2 \in B^2(P)$, $A \in \mathcal{B}(\mathbb{R}^3)$, $\alpha \in \mathbb{R}$ as:

$$(\nu_1 \oplus \nu_2)(A) =_{B(P)} \int_A \frac{d\nu_1}{dP} \frac{d\nu_2}{dP} dP, \quad (\alpha \odot \nu_1)(A) =_{B(P)} \int_A \left(\frac{d\nu_1}{dP} \right)^\alpha dP,$$

and becomes a separable Hilbert space when equipped with a proper inner product [4]. In the following subsection, we focus on the special case of Gaussian compartmental diffusion for which analytic expressions can be obtained for the perturbation and powering operations as well as for the distance.

2.2   Gaussian Compartmental Diffusion

The diffusion signal arising from Gaussian compartmental diffusion reads [2]:

$$S(\mathbf{q}) = S_0 e^{-\mathbf{q}^\top D \mathbf{q}}, \tag{3}$$

where D is the so-called diffusion tensor, i.e. a 3×3 symmetric positive definite (SPD) matrix. Hence, the probability measure ν induced by S is a Gaussian distribution with mean $\mathbf{0}$ and covariance matrix $D^{-1}/2$ and we can write, for the purpose of notation, $S \sim \mathcal{N}(\mathbf{0}, D^{-1}/2)$. As a result, Gaussian diffusion signals

cannot be embedded in $B^2(\lambda)$ because choosing the Lebesgue measure as reference makes the integral of the squared logarithm of S diverge. However since all moments of the Gaussian distribution are finite, if we define the reference measure P as a Gaussian distribution with mean $\mathbf{0}$ and covariance matrix D_{ref} SPD, i.e. $P \sim \mathcal{N}(\mathbf{0}, D_{\text{ref}})$, then Gaussian diffusion signals can be embedded in $B^2(P)$. Furthermore, it is possible to obtain analytic expression for the operations $\oplus$ and $\odot$. The perturbation of measure ν_1 by measure ν_2 reads:

$$(\nu_1 \oplus \nu_2)(A) =_B \int_A \frac{d\nu_1}{d\lambda} \frac{d\nu_2}{d\lambda} \frac{d\lambda}{dP} d\lambda =_B \int_A e^{-\mathbf{q}^\top (D_1 + D_2 - D_{\text{ref}})\mathbf{q}} d\lambda(\mathbf{q}).$$

Hence, $(\nu_1 \oplus \nu_2) \sim \mathcal{N}(\mathbf{0}, D_{\text{ref}} + (D_1 - D_{\text{ref}}) + (D_2 - D_{\text{ref}}))$. The result is another Gaussian distribution whose covariance matrix is centered around the one of the reference Gaussian distribution but perturbed by the covariance information in D_1 and D_2 that are not already present in the reference covariance structure. The multiplication of measure ν by a scalar $\alpha \in \mathbb{R}$ reads:

$$(\alpha \odot \nu)(A) =_B \int_A \left(\frac{d\nu}{d\lambda} \right)^\alpha \left(\frac{d\lambda}{dP} \right)^{\alpha - 1} d\lambda =_B \int_A e^{-\mathbf{q}^\top (\alpha D + (1-\alpha)D_{\text{ref}})\mathbf{q}} d\lambda(\mathbf{q}).$$

Hence, $(\alpha \odot \nu) \sim \mathcal{N}(\mathbf{0}, \alpha D + (1 - \alpha)D_{\text{ref}})$. The result is another Gaussian distribution whose covariance matrix is a linear combination between the covariances of ν and P. This offers the possibility of giving more or less weight to the information content of a diffusion signal in terms of microstructure. If one think of $\alpha \in [0, 1]$, then $(\alpha \odot \nu)$ means that less weight is put on the information content of ν (which could come from a poor model estimate, low SNR, etc) in favor of the information content of the reference P. In practice, it shrinks the diffusion tensor of the signal towards the one of the reference.

Note that writing the results of the operations of $\oplus$ and $\odot$ as Gaussian distributions is an abuse of notation. In effect, the space of diffusion signals is a subspace of $B^2(P)$ that is not closed in the Bayes geometry. This is because operations on diffusion signals in $B^2(P)$ might yield non positive "covariance" matrices. However, most applications in which analytic diffusion model computing is required involve weighted average operations, i.e. linear combinations with positive weights whose sum is less than one, which always produce SDP tensors in $B^2(P)$. In effect, it is easy to show that:

$$\bigoplus_{i=1}^N w_i \odot \nu_i \sim \mathcal{N}\left(\mathbf{0}, \sum_{i=1}^N w_i D_i + \left(1 - \sum_{i=1}^N w_i \right) D_{\text{ref}} \right). \qquad (4)$$

In the rare events that operations in Bayes space produce non-positive matrices, one can perform an orthogonal projection of the non-positive matrix back into the space of diffusion signals by solving the following minimization problem:

$$\min_{D \in \mathcal{B}(P) \text{ s.t. } D > 0} a d_{\mathcal{B}(P)}^2(M, D) + (1 - a)d_{\mathcal{B}(P)}^2(D_{\text{ref}}, D), \quad a \in [0, 1], \qquad (5)$$

where M is the symmetric matrix assumed to have some negative eigenvalues and $d_{\mathcal{B}(P)}$ is the distance on $B^2(P)$ given by:

$$\|\nu_1 \ominus \nu_2\|_{\mathcal{B}(P)}^2 := \text{Tr} \left((D_1 - D_2)D_{\text{ref}}^{-1}(D_1 - D_2)D_{\text{ref}}^{-1} \right) = \|(D_1 - D_2)D_{\text{ref}}^{-1}\|_F^2.$$

3   Experimental Setup

3.1   Simulations: Robustness to Noise

The goal of the simulated data is to assess the robustness of Bayes space interpolation to MRI-induced noise, compared to more traditional approaches that focus directly on the tensor and embed it either in Euclidean or log-Euclidean space. In Bayes space, as shown in Eq. (4), the tensor associated with the interpolated signal is a convex combination of the tensor interpolated in Euclidean space and the tensor of the reference signal, where the weight w_{data} associated with the tensor interpolated in Euclidean space shall indicate how much we are willing to trust its information content. For this purpose, we set $w_{\text{data}} := 1 - e^{-\text{SNR}/\beta}$ and include in the comparison three Bayes spaces (Bayes10, Bayes20, Bayes30), using $\beta = 10, 20, 30$ respectively. We generate (i) a set of 31 noise-free diffusion signals according to Eq. (3) with a baseline signal $S_0 = 1$ and 31 diffusion gradients $\mathbf{q}$ uniformly distributed on the hemisphere of radius $\sqrt{b}$, with $b = 1000\text{s/mm}^2$ and (ii) a set of $n = 8$ normalized weights to define spatial weights of the 8 neighbors. Next, for a given SNR and a given method, we obtain a Monte-Carlo estimate of the mean-square error (MSE) of the interpolated tensor by averaging the squared distances between the ground truth tensor and $R = 1000$ replicates of interpolated tensors from noisy neighboring tensors produced as follows:

(a) Add Rician noise to the noise-free signals using $\sigma = S_0/\text{SNR}$;

(b) Get a realistic noisy tensor field by estimating a diffusion tensor in each neighboring voxel via maximum likelihood estimation [11];

(c) Interpolate the tensor in the central voxel from the tensors in the 8 neighboring voxels to which we associate the initially simulated spatial weights;

(d) Compute distance between the interpolated and ground truth tensor.

We used 8 SNR $\in [6.25, 50]$ and three metrics to compare interpolations with ground truth: the angle between principal orientations (*direction* distance), the Euclidean distance between axial diffusivities (*diffusivity* distance) and the Euclidean distance between radial diffusivities (*radius* distance as it is often used as a proxy to measure axon radii). We chose these metrics as they focus on microstructural parameters of direct clinical relevance and are independent from the compared spaces. For interpolation in Bayes space, for each SNR, we set the reference tensor as a noisy version of the ground truth tensor using the same procedure as above using $\text{SNR}\sqrt{N}$. This is because, often, the available data to use as reference comes from atlases generated on similar data but averaged on multiple subjects (N). Hence, there is still uncertainty in the reference tensor but its variance is likely to be divided by N. In this simulation, we set $N = 20$.

3.2   Case Study: Tractography of the Pyramidal Tract

The pyramidal tract (PT) is of primary importance as it handles volontary motion. Neurons initiate in the primary motor cortex (R0), goes successively through the corona radiata (R1), genu and posterior limb of the internal capsule (R2) and cerebral peduncles (R3) to eventually enter the spinal cord (R4). The

PT is difficult to reconstruct from tractography due to the large spreading of neurons on the cortex. We aim at showing that interpolation in Bayes space makes deterministic FACT tractography [7] feasible for PT reconstruction. In details, we compare four approaches obtained combining single- or multi-tensor FACT tractography (ST or MT) with log-Euclidean or Bayes interpolation (ST-LogEuclidean, ST-Bayes, MT-LogEuclidean and MT-Bayes).

We scanned a healthy subject on a 3T Siemens Verio magnet for a T_1 MPRAGE image at 1mm^3 resolution and diffusion data at 2mm^3 resolution using the same gradient table as in the simulations. We used an available PT atlas in MNI space based on 20 healthy subjects that underwent the same acquisition protocol. We estimated ST and MT models from the diffusion data [11] and brought the resulting images into MNI space where FACT tractography was performed. In the case of MT-FACT, at a given point, a single tensor was picked from each neighboring mixture according to the highest orientation similarity w.r.t. arrival direction so that we could proceed as in ST-FACT. We performed tractography of the PT following regions of interest (ROI). We used R0 as seeding mask, stopped a streamline generation when FA fell below 0.1 (for ST-FACT) or when linearly interpolated fraction of free water exceeded 0.8 (for MT-FACT) and filtered the resulting streamlines progressively through R1, R2, R3 and R4. We defined the reference weight at position $\mathbf{x}$ for Bayes interpolation as:

$$w_{\mathrm{ref}}(\mathbf{x}, \mathrm{FA}, \mathrm{SNR}, \mathbf{x}_{\mathrm{ref}}; \beta, \delta) := \left(1 - \mathrm{FA}(1 - e^{-\mathrm{SNR}/\beta})\right) e^{-\|\mathbf{x} - \mathbf{x}_{\mathrm{ref}}\|/\delta},$$

where FA is the fractional anisotropy of the tensor interpolated in Euclidean space, $\mathbf{x}_{\mathrm{ref}}$ is the position of the closest point on the PT atlas and (β, δ) are user-defined parameters that control the decay velocity of SNR and distance weights. In essence, we put more trust in the data when both SNR and orientation coherence between neighbors are high or when the reference is too far.

4 Results & Discussion

Simulations: Robustness to Noise. Figure 1 shows the MSE between the ground truth tensor in central voxel and the interpolated tensor from noisy neighbors, with increasing amount of noise. For all metrics, interpolation in Bayes space performs uniformly better in terms of robustness to MRI-induced noise for recovering microstructural parameters. The MSE curves for the three Bayes spaces that give increasing importance to the reference measure (from Bayes10 to Bayes30) reveals that even when we trust mostly the data (Bayes10), interpolation in Bayes space is preferable.

Case Study: Tractography of the Pyramidal Tract. Figure 2 shows PT reconstruction with increasing number of filtering ROIs (from left to right) for all four methods. First, observe that only methods based on Bayes interpolation (rows 2 and 4) successfully manage to reconstruct streamlines that go through all four ROIs and thus are more likely to belong to the PT. Also, in general, deterministic FACT tractography fails to recover PT streamlines with the
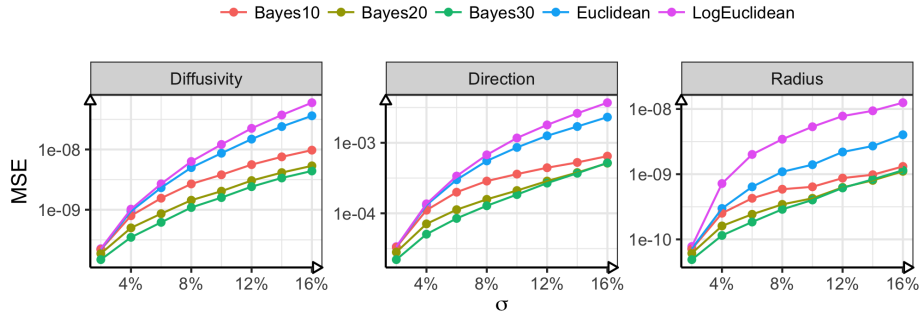
Fig. 1. MSE of interpolators as a function of noise. Axial diffusivity recovery (left), principal eigenvector recovery (middle), radial diffusivity recovery (right). Ground truth tensor: orientation $(\sqrt{2}/2, \sqrt{2}/2, 0)^\top$; diffusivities $10^{-3}(1.71, 0.3, 0.1)^\top \text{mm}^2/\text{s}$.
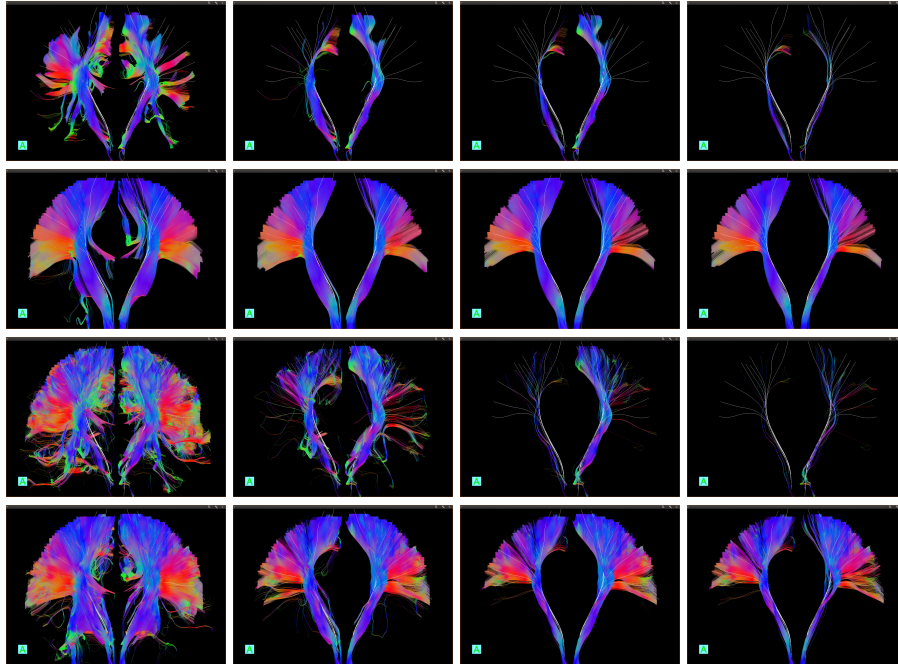
Fig. 2. Coronal View of Reconstructed Pyramidal Tracts (with overlaid PT atlas). Columns correspond to increasing number of filtering ROIs: R1 (1st column), R1+R2 (2nd column), R1+R2+R3 (3rd column), R1+R2+R3+R4 (4th column). Rows correspond to the methods (from top to bottom): single-tensor FACT – log-Euclidean, and Bayes, multi-tensor FACT – log-Euclidean, and Bayes.

traditional log-Euclidean interpolation (rows 1 and 3) . This is well documented in the literature on ST-FACT. We hypothesize that MT-FACT has a lot more

directional information for tracking but, without prior anatomical knowledge, the FACT algorithm (which stepwise follows the most collinear direction) has no mechanism to channel streamlines into following the PT shape. This explanation is supported by the sequential filtering that shows high streamline variability when filtering only by the corona radiata (1st column, 3rd row) and almost no remaining streamlines at the end of the entire filtering process (4th column, 3rd row). The ST-Bayes method seems to mainly follow the shape of atlas streamlines. This is due to an inherent model mis-specification in most parts of the white matter where the ST model provides an insufficient description of the microstructure and therefore presents an articially low FA which uniformly inflates the weight of the reference measure. Conversely, the MT-Bayes version nicely preserves the PT streamlines after complete filtering without heavy influence of the atlas since the MT model provides an accurate description of the microstructure and thus the interpolated tensor has low FA only when neighboring tensors have heterogeneous orientations.

References

1. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. MRM 56(2), 411–21 (2006)
2. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson. B 103(3), 247–54 (1994)
3. van den Boogaart, K.G., Egozcue, J.J., Pawlowsky-Glahn, V.: Bayes Hilbert Spaces. Aust. N. Z. J. Stat. 56(2), 171–194 (2014)
4. Egozcue, J.J., Pawlowsky-Glahn, V., Tolosana-Delgado, R., Ortego, M.I., van den Boogaart, K.G.: Bayes spaces: use of improper distributions and exponential families. Rev. Real Acad. Ciencias Exactas, Fis. Nat. Ser. A Mat. 107(2), 475–86 (2013)
5. Le Bihan, D.: Looking into the functional architecture of the brain with diffusion MRI. Nat. Rev. Neurosci. 4(6), 469–80 (2003)
6. Lenglet, C., et al.: Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. JMIV 25(3), 423–44 (2006)
7. Mori, S., Crain, B., et al.: Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Annals of neurology 45(2), 265–9 (1999)
8. Panagiotaki, E., et al.: Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. Neuroimage 59(3), 2241–54 (2012)
9. Pennec, X., Fillard, P., Ayache, N.: A riemannian framework for tensor computing. International Journal of Computer Vision 66(1), 41–66 (2006)
10. Scherrer, B., Warfield, S.K.: Parametric representation of multiple white matter fascicles from cube and sphere diffusion MRI. PLoS One 7(11), e48232 (2012)
11. Stamm, A., Commowick, O., et al.: Comprehensive Maximum Likelihood Estimation of Diffusion Compartment Models Towards Reliable Mapping of Brain Microstructure. In: MICCAI (III). LNCS, vol. 9902, pp. 622–630 (2016)
12. Taquet, M., Scherrer, B., et al.: A mathematical framework for the registration and analysis of multi-fascicle models for population studies of the brain microstructure. IEEE TMI 33(2), 504–17 (2014)
13. Wang, et al.: When Brownian diffusion is not Gaussian. Nat Mater 11(6), 481–5 (2012)