

Learning a Reliable Estimate of the Number of Fiber Directions in Diffusion MRI

Thomas Schultz

Max Planck Institute for Intelligent Systems, Tübingen, Germany*

Abstract. Having to determine an adequate number of fiber directions is a fundamental limitation of multi-compartment models in diffusion MRI. This paper proposes a novel strategy to approach this problem, based on simulating data that closely follows the characteristics of the measured data. This provides the ground truth required to determine the number of directions that optimizes a formal measure of accuracy, while allowing us to transfer the result to real data by support vector regression. The method is shown to result in plausible and reproducible decisions on three repeated scans of the same subject. When combined with the ball-and-stick model, it produces directional estimates comparable to constrained spherical deconvolution, but with significantly smaller variance between re-scans, and at a reduced computational cost.

1 Introduction

Multi-compartment models are a traditional way of estimating more than a single fiber orientation in diffusion MRI [1,2]. The number of fiber compartments used in such models can have a profound effect on the estimated directions, making it mandatory to decide on a setting that is adequate for any given voxel.

Despite this, only few systematic approaches to this problem are available: Automated Relevance Determination [2] has been used to force the weights of fiber compartments with insufficient statistical support to zero. However, it requires computation in a full Bayesian framework. The Bayesian Information Criterion has been demonstrated to produce suboptimal results even on idealized synthetic data [3]. Approaches based on peaks of the fiber ODF [3,4] require setting a threshold, and suffer from noise-induced spurious peaks.

This work proposes a novel strategy for setting the number of fiber compartments. It formalizes the pragmatic view that it is best to use the number that leads to the most accurate estimates of the desired parameters. Since the in vivo data lacks ground truth, the optimal number of compartments is determined in simulated data that closely follows the characteristics of the experimental data. Based on this synthetic data, a classifier is trained to recognize voxels that are best analyzed by a ball-and-stick model [2] with a single, two, or three fiber compartments, and thereby achieves reliable and efficient fiber direction estimates.

* I would like to thank Lek-Heng Lim (University of Chicago) for useful discussions and Hans J. Johnson (University of Iowa) and the PREDICT-HD project for sharing the data set (made possible by NIH grants NS054893, U54EB005149, NS40068).

2 Defining the “Most Adequate” Number of Directions

When analyzing a diffusion MRI signal that arises as an average over n fiber compartments with principal directions \mathbf{v}_i and relative weights w_i , two or more of the \mathbf{v}_i may be so close to collinearity that, given noisy measurements with limited angular and spectral resolution, it becomes impossible to separate them with reasonable precision. In this case, we prefer to describe them with a single estimate $\hat{\mathbf{v}}_j$. Similarly, when some of the weights w_i are so small that the associated directions can no longer be reliably estimated, we prefer to set them to zero in order to obtain a less complex model that is more robust to noise.

This tradeoff is formalized by the following definition of weighted average angular error (WAAE), which measures the average angle between each true direction \mathbf{v}_i and its nearest estimate $\hat{\mathbf{v}}_j$, weighted by its true volume fraction w_i :

$$\text{WAAE} := \sum_{i=1}^n w_i \min_j \arccos(|\mathbf{v}_i \cdot \hat{\mathbf{v}}_j|) \quad (1)$$

Given a model that requires choosing the number of fiber compartments, we define the number that minimizes WAAE as the one most suitable for analysis. Even though WAAE is used throughout this paper, the fundamental idea is to use some formal error measure like it to decide on the “right” number of directions, not necessarily its exact definition. In particular, when dealing with advanced multi-compartment models that additionally estimate parameters such as axon diameter [5] or account for fiber fanning or bending, learning could be based on an objective function that penalizes errors in those.

3 Learning the Number of Fiber Compartments

Since evaluation of WAAE requires ground truth, it cannot be computed directly for in vivo data. Instead, we use machine learning to train a classifier that predicts the best number of fiber compartments for experimental data based on its similarity to simulated data, for which WAAE can be computed.

3.1 Support Vector Regression

Given a suitable representation \mathbf{x} of the diffusion-weighted signal $S(\theta, \phi)$, we seek a function $\hat{f}(\mathbf{x})$ whose value provides an estimate of the most adequate number of directions to analyze S . We obtain $\hat{f}(\mathbf{x})$ through support vector regression, which produces functions of the form

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) k(\mathbf{x}_i, \mathbf{x}) + b. \quad (2)$$

Support vector regression requires us to decide on a representation \mathbf{x} of the signal, to specify training samples \mathbf{x}_j for which the value of $f(\mathbf{x}_j)$ is known,

and to specify a kernel function $k(\mathbf{x}_j, \mathbf{x})$ that measures the similarity between any input \mathbf{x} and the training data \mathbf{x}_j . Based on this, an optimization procedure automatically finds a subset $\{\mathbf{x}_i\}_{i=1}^{\ell}$ of the training data and suitable parameters α_i , α_i^* , and b , so that a given distance measure between $\hat{f}(\mathbf{x}_j)$ and the training values $f(\mathbf{x}_j)$ is minimized, subject to additional constraints. Full details of the method are beyond the scope of this paper and are given in [6].

3.2 Feature Definition and Kernel Selection

The representation \mathbf{x} from which Eq. (2) is computed is known as a feature vector. Ideally, it should encode the available prior knowledge about f . For example, the number of fiber compartments clearly should not be affected by joint rotations of all fibers. Therefore, \mathbf{x} should be invariant under rotations of the measurement frame, while providing information that can be used to infer a suitable number of directions.

The experiments reported in this paper use the three sorted eigenvalues of a single diffusion tensor per voxel as the feature vector. Even though the diffusion tensor is insufficient to resolve the *directions* of more than a single compartment, the results show that it reliably indicates their *number*. Experiments with more complex features based on rotational invariants of various HARDI models, including ADC profiles [7], Q-Balls, and fiber ODFs, have led to small improvements on simulated data, but reproducibility on real data was much reduced.

As the kernel $k(\mathbf{x}_j, \mathbf{x})$, we use a standard radial basis function, $k(\mathbf{x}_j, \mathbf{x}) = \exp(-\gamma\|\mathbf{x}_j - \mathbf{x}\|^2)$. The parameter γ , as well as two additional parameters, C and ν , which control the number ℓ of support vectors in Eq. (2), have been fixed automatically using cross-validation [6].

3.3 Training Data and Labels

In order to successfully transfer the learned function $\hat{f}(\mathbf{x})$ to real data, it is essential that the characteristics of the training data \mathbf{x}_j be as similar as possible to the experimental data. Therefore, we generate it directly from the measurements, with random weights $w_i > 0$, $w_1 + w_2 + w_3 = 1$, and random directions \mathbf{v}_i . Note that even though we always simulate three compartments, modeling them with only one or two compartments may lead to a smaller WAAE when some w_i are very small, or at least two \mathbf{v}_i are nearly collinear.

As in [8], the N voxels with largest FA in the measured data are assumed to contain a single dominant direction, indicated by the principal diffusion direction. It is our goal to reproduce the natural variability of these single fiber voxels, including the bending and spreading which is present even in the most anisotropic voxels. Therefore, we randomly select one of the N voxels as a template for each simulated fiber compartment. In order to get a sufficiently realistic estimate of the variability, we use $N = 1000$, corresponding to $\text{FA}_{\min} \approx 0.68$.

Since the classifier should be able to handle partial voluming with cerebrospinal fluid (CSF) or gray matter, we include it in our training data. For this, we use the N *least* anisotropic voxels each from a CSF mask, obtained by Otsu

segmentation [9] on mean diffusivity, and a brain mask excluding CSF, respectively. Each simulated voxel \mathcal{V} selects a near-isotropic signal $I(\theta, \phi)$ from either gray matter or CSF, and weights it by a uniform random factor $w_{\text{iso}} \in [0, 0.5]$.

Taken together, the simulated signal S is given by the following weighted combination of a randomly selected near-isotropic signal I , three randomly selected single-fiber signals R_i , rotated to match the directions \mathbf{v}_i , and error terms ε that are estimated using residual bootstrapping [10]:

$$S(\theta, \phi) = w_{\text{iso}}(I(\theta, \phi) + \varepsilon_{\text{iso}}) + (1 - w_{\text{iso}}) \sum_{i=1}^3 w_i (R_i(\theta, \phi) + \varepsilon_i) \quad (3)$$

From the simulated data, eigenvalue features \mathbf{x}_j are computed, and directional estimates $\hat{\mathbf{v}}_j$ are obtained from the ball-and-stick model [2] with one, two, and three fiber directions. The number of compartments that led to the smallest WAAE is used as the true value of $f(\mathbf{x}_j)$ in the support vector regression. No cases other than $\{1, 2, 3\}$ are considered, since we assume a separate classifier defines a white matter mask, and no prior work has plausibly reconstructed more than three fiber directions from a single voxel.

Since we only use discrete values $f(\mathbf{x}_j) \in \{1, 2, 3\}$ in our training data, we should avoid including boundary cases in which two settings perform similarly well. Therefore, we simulate a large number of voxels (250,000) and only train on the 1,000 examples \mathbf{x}_j of each class for which the difference in WAAE between the optimal choice of compartment number and the second-best choice was largest.

4 Results on Experimental Data

Three repeated diffusion MR acquisitions have been obtained within the same session at 3 T, with voxel size $2 \times 2 \times 2 \text{ mm}^2$, 71 gradient directions, 8 B0 images, $b = 1000 \text{ s/mm}^2$. Eddy current distortions and head motion have been corrected for using FSL (www.fmrib.ox.ac.uk), and the B matrix has been rotated accordingly [11]. In order to evaluate reproducibility, results that use all available data are compared to results obtained from the three individual repeats.

4.1 Number of Fibers

We clamp the values of $\hat{f}(\mathbf{x})$ to range $[1, 3]$ and round them to obtain discrete classes $\{1, 2, 3\}$. The percentages of voxels that were marked as being best analyzed with one, two, or three fiber compartments are listed in Table 1 for a brain mask (excluding CSF), and for two different thresholds of Fractional Anisotropy.

A comparison to constrained spherical deconvolution (with $l_{\text{max}} = 8$) was performed using the software MRtrix [8]. Applying an FOD threshold of 0.1 as in [4] results in a much smaller number of single fiber voxels than with our classifier. Moreover, compared to the combined data from all three repeats, deconvolution systematically reported a smaller number of single-fiber voxels in the individual repeats (between 2.3% and 2.4%, rather than 3.4%).

Table 1. Counting the peaks in constrained deconvolution (CSD) estimates fewer single-fiber voxels than our classifier, consistent with known effects of noise on CSD

	Learned Classifier			Non-Neg. Deconvolution		
	non CSF	FA > 0.1	FA > 0.2	non CSF	FA > 0.1	FA > 0.2
1 Fiber	7.6 %	9.8 %	15.8 %	3.4 %	4.4 %	7.3 %
2 Fibers	20.3 %	26.1 %	43.6 %	18.6 %	24.1 %	40.6 %
3 Fibers	72.1 %	64.1 %	40.6 %	78.0 %	71.5 %	52.1 %

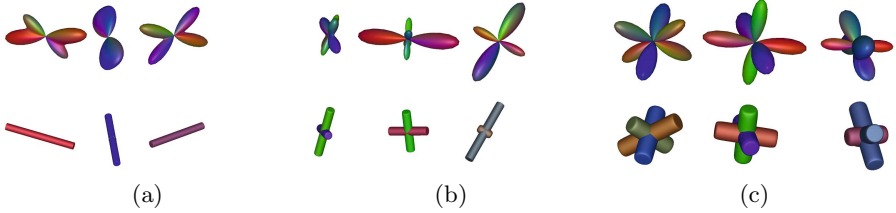


Fig. 1. While fiber ODFs from constrained deconvolution (top row) show reasonable agreement with the ground truth (bottom row) in simulated two- and three-fiber cases (b/c), large spurious peaks arise in two out of the three shown single directions (a)

This indicates that the reduced effective SNR in the individual measurements leads to spurious peaks in the fiber ODFs, which reduces the reliability of their number as an indicator of distinct fiber compartments. Figure 1 illustrates the problem using three examples each from the training sets that define the single-, two-, and three-fiber cases (a–c). The rods in the bottom row indicate the ground truth directions \mathbf{v}_i , lengths indicating relative weights w_i (radii being reduced in (a) and (b) to avoid occluding the smaller contributions in (b)).

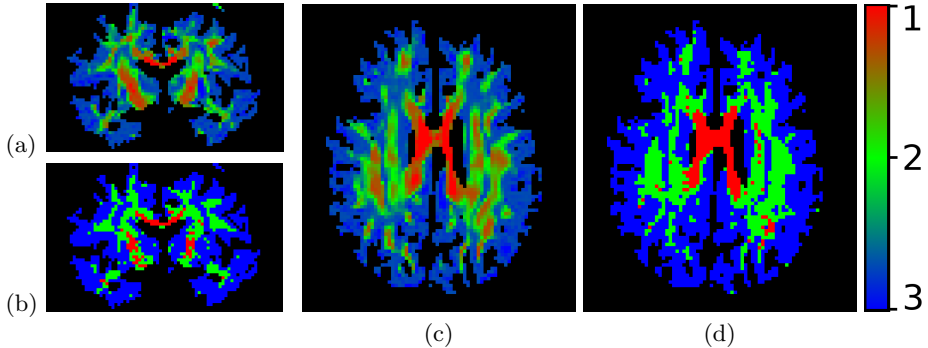
The strong spurious peaks that arise in two out of the three single-fiber cases (a) are a known problem of deconvolution, particularly in the presence of partial voluming with isotropic compartments [12]. They are caused by the fact that the regularization introduced in [8] suppresses negative peaks, but does not address spurious positive peaks. In contrast to this, the fraction of fibers assigned to each class by our classifier was stable across re-scans. The percentage of single-fiber voxels in the individual repeats varied between 7.5% and 7.9%.

The reproducibility of our estimates is further confirmed by Table 2, which lists the percentage of voxels in which the classification in each individual measurement agreed with the class assigned based on the combined data, as well as the mean difference and the 95% confidence interval of the value of $\hat{f}(\mathbf{x})$.

Figure 2 presents a visual comparison between our classification (a/c) and the results from fODF thresholding (b/d) in a coronal (a/b) and an axial (c/d) slice. As in [4], red, green, and blue indicate one, two, and three directions, respectively. In (a/c), the values of $\hat{f}(\mathbf{x})$ are mapped before rounding, to demonstrate the smooth transition between the classes.

Table 2. Good agreement between classification of the combined data and individual repeats supports the reproducibility of our method

	Learned Classifier			Non-Neg. Deconvolution		
	Agreement	Mean	95 % Conf.	Agreement	Mean	95 % Conf.
1st Repeat	94.5 %	0.062	0.215	89.7 %	0.106	1.0
2nd Repeat	95.3 %	0.056	0.189	90.0 %	0.103	1.0
3rd Repeat	95.3 %	0.056	0.192	90.1 %	0.102	1.0

**Fig. 2.** Compared to the fiber number estimate from constrained deconvolution (b/d), our classifier (a/c) provides more coherent clusters of single fiber voxels (red), and smooth transitions between the classes

4.2 Estimates of Fiber Direction

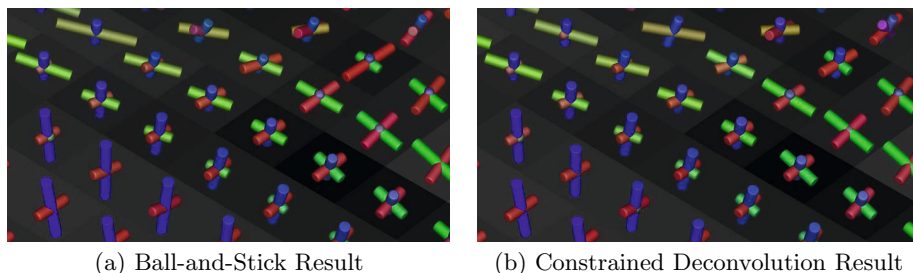
Since the parameter we are ultimately interested in is fiber orientation, let us now consider the reproducibility of directional estimates based on the ball-and-stick model [2] when the number of sticks is determined by our function $\hat{f}(\mathbf{x})$.

Table 3 reports the weighted average angular deviation (WAAD) within a white matter mask ($FA > 0.2$). It is computed from Eq. (1), by treating the estimates from the combined data as “ground truth”. The results from all three repeats were very similar and have been averaged for presentation. Compared to constrained deconvolution, ball-and-stick achieves slightly lower precision in the two-fiber case, but higher reproducibility in the one- and three-fiber cases. The unfavorable 90% confidence bound on deconvolution-based single fiber estimates is consistent with the emergence of spurious peaks as observed in Figure 1 (a).

The mean WAAD over the whole white matter was 9.8° when combining ball-and-stick with our classifier, which improves over the WAAD achieved by deconvolution (11.0°). According to a two-sided t-test on the distribution of WAADs from all white matter voxels, this difference is highly significant ($p < 10^{-20}$), in each of the three repeated measurements. In contrast, fitting the ball-and-stick model with three compartments in all voxels led to a larger average WAAD (11.9°). This confirms the importance of selecting an adequate number of fibers based on the data.

Table 3. Reasonable agreement is achieved between direction estimates from the combined data and individual measurements for one- and many two-fiber voxels

	Ball-and-Stick			Non-Neg. Deconvolution		
	Mean	Median	90 % Conf.	Mean	Median	90 % Conf.
1 Fiber	2.0	1.2	3.0	4.0	1.9	10.5
2 Fibers	8.7	5.7	20.9	7.1	5.4	14.7
3 Fibers	14.1	13.5	23.6	15.0	14.6	24.6

**Fig. 3.** The triple crossing between corpus callosum (red), corticospinal tract (blue) and superior longitudinal fasciculus (green) was the only region that produced reproducible three-fiber estimates. Closeup shows a slanted view onto an axial slice.

Large absolute errors indicate that many of the voxels that have been labeled as “three-fiber” by both methods do not afford reliable directional estimates using either model. However, inspecting all voxels in which three-fiber estimates were consistently obtained with less than 10° WAAD revealed two clear clusters: The triple crossing between corpus callosum, corticospinal tract, and superior longitudinal fasciculus, in both hemispheres. Visual inspection of a detail of that region in Figure 3 suggests that, when combined with an appropriate classifier, ball-and-stick fitting produces directional estimates that are very similar to constrained deconvolution.

The proposed method is computationally efficient. It took less than 10 seconds to propose a fiber number for all 96,000 voxels within the brain mask. Subsequent fitting of the ball-and-stick model took 74 seconds. In comparison, constrained deconvolution and subsequent peak finding with the implementation from [8] took almost 10 minutes on the same 2.7 GHz workstation.

5 Conclusion

A novel strategy to select the number of fiber directions in multi-compartment models has been presented, which explicitly aims to minimize a formal measure of error in the estimated model parameters. Since accuracy cannot be measured in vivo, a machine learning approach is used to automatically transfer insights from simulated data with known ground truth to experimental data.

On the conceptual side, the main contribution of this work is to demonstrate that such a transfer produces plausible and reproducible results. As a practical benefit, combining the proposed classifier with the ball-and-stick model produces directional estimates that are similar overall, but more reproducible than the ones from constrained deconvolution across re-scans, particularly in single-fiber voxels. They are also obtained at a markedly reduced computational cost. This seems particularly relevant when performing bootstrapping-based tractography, which requires repeated model fitting [13].

References

1. Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine* 48, 577–582 (2002)
2. Behrens, T.E.J., Johansen-Berg, H., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* 34, 144–155 (2007)
3. Schultz, T., Westin, C.-F., Kindlmann, G.: Multi-Diffusion-Tensor Fitting via Spherical Deconvolution: A Unifying Framework. In: Jiang, T., Navab, N., Plum, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010, Part I. LNCS*, vol. 6361, pp. 674–681. Springer, Heidelberg (2010)
4. Jeurissen, B., Leemans, A., Tournier, J.D., Jones, D.K., Sijbers, J.: Estimating the number of fiber orientations in diffusion MRI voxels: a spherical deconvolution study. In: *Proc. Int'l Society of Magnetic Resonance in Medicine, ISMRM* (2010)
5. Zhang, H., Dyrby, T.B., Alexander, D.C.: Axon Diameter Mapping in Crossing Fibers with Diffusion MRI. In: Fichtinger, G., Martel, A., Peters, T. (eds.) *MICCAI 2011, Part II. LNCS*, vol. 6892, pp. 82–89. Springer, Heidelberg (2011)
6. Schölkopf, B., Smola, A.J.: *Learning with Kernels*. MIT Press (2002)
7. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine* 48, 331–340 (2002)
8. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* 35, 1459–1472 (2007)
9. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Trans. on Systems, Man and Cybernetics* 9(1), 62–66 (1979)
10. Chung, S., Lu, Y., Henry, R.G.: Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. *NeuroImage* 33(2), 531–541 (2006)
11. Leemans, A., Jones, D.K.: The B-matrix must be rotated when correcting for subject motion in DTI data. *Magnetic Resonance in Medicine* 61, 1336–1349 (2009)
12. Dell'Acqua, F., Scifo, P., Rizzo, G., Catani, M., Simmons, A., Scotti, G., Fazio, F.: A modified damped richardson-lucy algorithm to reduce isotropic background effects in spherical deconvolution. *NeuroImage* 49, 1446–1458 (2010)
13. Jeurissen, B., Leemans, A., Jones, D.K., Tournier, J.D., Sijbers, J.: Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. *Human Brain Mapping* 32, 461–479 (2011)