The Importance of Being Dispersed: A Ranking of Diffusion MRI Models for Fibre Dispersion Using *In Vivo* Human Brain Data

Uran Ferizi[1,2], Torben Schneider[2], Maira Tariq[1], Claudia A.M. Wheeler-Kingshott[2], Hui Zhang[1], and Daniel C. Alexander[1]

[1] CMIC, Dept. Computer Science and Dept. Medical Physics and Bioengineering, University College London, United Kingdom
[2] NMR Research Unit, Department of Neuroinflammation, Institute of Neurology, University College London, United Kingdom
uran.ferizi.10@ucl.ac.uk

Abstract. In this work we compare parametric diffusion MRI models which explicitly seek to explain fibre dispersion in nervous tissue. These models aim at providing more specific biomarkers of disease by disentangling these structural contributions to the signal. Some models are drawn from recent work in the field; others have been constructed from combinations of existing compartments that aim to capture both intracellular and extracellular diffusion. To test these models we use a rich dataset acquired *in vivo* on the corpus callosum of a human brain, and then compare the models via the Bayesian Information Criteria. We test this ranking via bootstrapping on the data sets, and cross-validate across unseen parts of the protocol. We find that models that capture fibre dispersion are preferred. The results show the importance of modelling dispersion, even in apparently coherent fibres.

1 Introduction

Diffusion MRI probes the tissue microstructure, by measuring the water dispersion in biological tissue. This technique is often applied in the brain, especially where parallel fibres restrict the water mobility anisotropically, thus providing putative measures of white matter integrity and connectivity.

Currently, the standard model for imaging diffusion in tissue is the diffusion tensor (DT) [1], which assumes a trivariate Gaussian dispersion pattern. Derived indices, e.g. mean diffusivity or fractional anisotropy, can correlate with major tissue damage, but lack the sensitivity and the specificity to provide indices such as axon radius, density, orientation and permeability. Stanisz et al. [2] pioneered a multi-compartment representation of separate diffusive processes in nervous tissue. The Ball-and-Stick model, by Behrens et al. [3] is the simplest possible two-compartment model with restricted axonal diffusion and isotropic extra-axonal diffusion. A recent class of parametric models has emerged to describe data better by additionally accounting for fibre directional incoherence, which is abundant in the brain, even at a sub-voxel level.

K. Mori et al. (Eds.): MICCAI 2013, Part I, LNCS 8149, pp. 74–81, 2013.

Ball-and-Sticks [3] can have more-than-one intracellular diffusion compartments. Zhang et al. [4] constructed NODDI to describe fibres with an explicit orientation dispersion index derived from a Watson distribution and tests the model with *in-vivo* human whole-brain data. Sotiropoulos et al. [5] design Ball-and-Rackets to describe fibre fanning through a Bingham distribution by extending the Ball-and-Sticks model [3]. The Bingham distribution extends the Watson distribution to account for asymmetric/anisotropic dispersion. This model is then applied to post-mortem macaque monkey brain data.

In this work, similar to the taxonomy provided in Panagiotaki et al. [6], we construct models that combine Ball (for isotropic diffusion), Zeppelin (for 2D anisotropic diffusion) or Tensor (for 3D anisotropic diffusion) for extracellular diffusion, with various models for intracellular diffusion: two-Sticks, a Watson or Bingham distribution of Sticks. We also add a further compartment for isotropically restricted diffusion: Dot (a zero radius sphere) or CSF. We then fit these models to a very rich dataset and, in addition to the fitting quality, we take into account the model complexity by using the Bayesian Information Criterion in order to discover which models explain the data best. Lastly, to validate the *BIC* ranking, we test the models through both bootstrapping on the data sets and prediction of unseen parts of the protocol.

2 Methods

This section first describes the models, then the data acquisition and the preprocessing done to obtain a set of measurements for fitting the models. Lastly, we detail the fitting procedure and the criterion applied to compare the models.

2.1 Models

Generic Model: The signal for a model with two or more types of compartments can be expressed as: $S = S_0 \left\{ \sum f_{ic}^k S_{ic}^k + f_{rc} S_{rc} + (1 - f_{rc} - \sum f_{ic}^k) S_{ec} \right\}$, where f_{ic} is the weight of the intracellular signal compartment S^{ic}, f_{rc} is the weight of the isotropically restricted signal compartment S_{rc}, S_{ec} is the extracellular signal compartment, and k is the compartment index.

Extracellular Compartments: The compartments used to capture signal outside the axons and the isotropically restricted compartments are the Tensor, the Zeppelin, the Zeppelin with tortuosity and the Ball. *Tensor* signal is modelled through the DT, as $S = \exp[\lambda_1(\mathbf{e_1 e_1^t}) + \lambda_2(\mathbf{e_2 e_2^t}) + \lambda_3(\mathbf{e_3 e_3^t})]$, where $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are the three characteristic vectors that define the 3D orientation of the Tensor and, along those directions, λ_1, λ_2 and λ_3 give the size of the tensor (or "apparent diffusivity"). The *Zeppelin* is a special case of the Tensor where $\lambda_2{=}\lambda_3$. We follow Szafer et al. [7] to express the *Zeppelin with tortuosity*, which has the radial diffusivity expressed in terms of the axial one, $\lambda_2{=}\lambda_1 f_{ic}$. Further, setting all three eigenvalues the same makes a *Ball*, where $\lambda_1{=}\lambda_2{=}\lambda_3$.

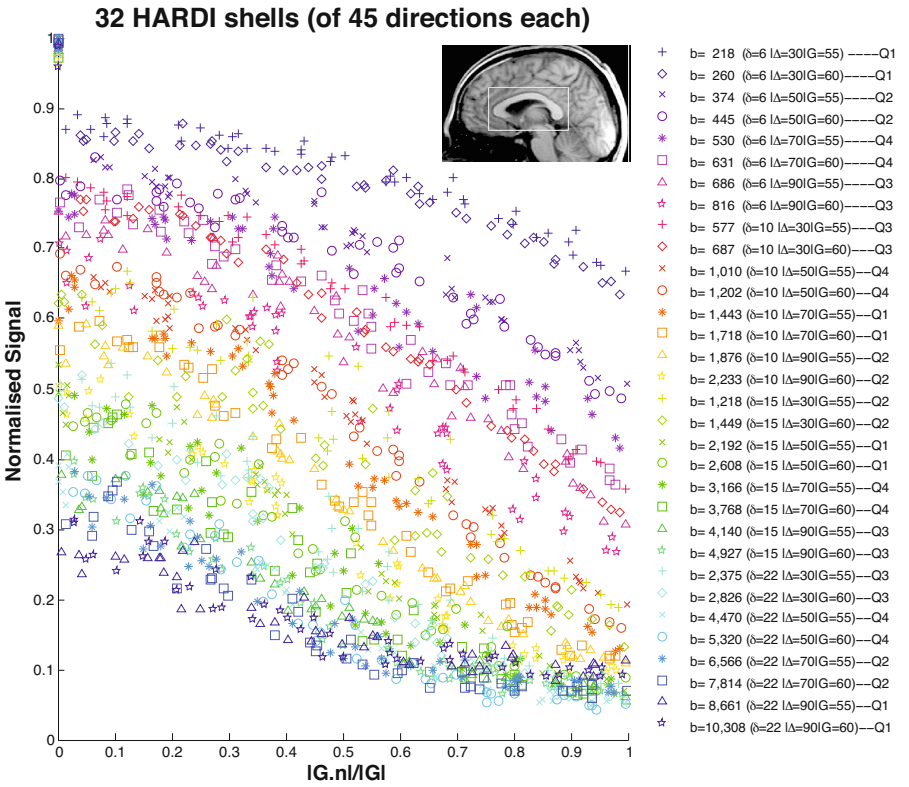

Fig. 1. The 2° data set. The legend gives b-value ($\delta \mid \Delta \mid |\mathbf{G}|$) in units of s/mm^2(ms|ms|mT/m); Q1-Q4 on the right define the four quarters of the full protocol used in the cross-validation. The insert picture is of a sagittal slice of the brain; boxed is the scanned volume, encompassing the corpus callosum. **G** is the applied gradient vector and **n** is the fibre direction; the x-axis gives the absolute value of the cosine of the angle between the applied gradient and fibre direction: to the left, the gradient is perpendicular to the fibres; to the right, parallel.

Intracellular Compartments: Sticks are used to represent the axonal diffusion, via either a discrete set of Sticks [3] (we pick two) or an underlying Bingham/Watson fibre orientation distribution [4,5]. The *Bingham* distribution is $f(\mathbf{n}|\kappa_1, \kappa_2, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = [{}_1F_1(\frac{1}{2}, \frac{3}{2}, \kappa_1, \kappa_2)]^{-1}\exp[\kappa_1(\boldsymbol{\mu}_1 \cdot \mathbf{n})^2 + \kappa_2(\boldsymbol{\mu}_2 \cdot \mathbf{n})^2]$, where κ_1 and κ_2 are the concentration parameters, such that $\kappa_1 \geq \kappa_2 \geq 0$; the mutually orthogonal vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ indicate the orientation axes of fibre dispersion. This is similar to a bivariate Gaussian distribution with elliptical contours on the sphere. The denominator, ${}_1F_1$, is a confluent hypergeometric function of first kind [8]. The *Watson* distribution is a special case of the Bingham distribution, where there is only one κ and $\boldsymbol{\mu}$ ($\kappa_2 = 0$); this corresponds to circular contours on the sphere.

2.2 Data Acquisition and Preprocessing

We use a PGSE sequence on a 3T Phillips scanner, with cardiac gating and TR=4s. The full protocol uses 32 HARDI shells. Each shell has a unique set of 45 directions, randomly rotated to enhance the angular resolution. The protocol has a wide range of achievable b-values, 218 to 10,308 s/mm^2, combining δ = {6, 10, 15, 22}ms, Δ = {30, 50, 70, 90}ms, $|\mathbf{G}|$ = {55, 60}mT/m, and three interwoven b=0 acquisitions.

The data is acquired in two separate non-stop sessions, each lasting about 4.5hrs (the "2x4hr" dataset). The field-of-view is centred on the mid-sagittal slice of the corpus callosum (CC), where we assume coherently oriented CC fibres are perpendicular to the image plane. There are nine 4mm-thick sagittal slices, the image size is 64 x 64 and the in-plane resolution is 2mm x 2mm.

After manually registering all DW images to the unweighted image of the b=1,202s/mm^2 shell, using only image translations, we fit the DT to this b=1,202 shell to select a set of voxels with coherently oriented fibres. Voxels with FA>0.6 and principal eigenvector within η=2° of the assumed fibre direction are retained. There are 24 voxels remaining, all in 2 slices close to the mid-sagittal plane, and mostly in the genu. A similar procedure is performed with η=5°, which leaves 66 voxels, and deviation 10° which leaves 99 voxels, in both cases across the genu and mid-body.

Before fitting the models, the signal of each DW slice is normalised by the b=0 images with the same echo time (TE). A single dataset is then created by averaging the voxels selected above. Fig.1 shows the signal from the 2° dataset, containing 1,536=32*(3+45) measurements.

2.3 Model Fitting and Selection

We use the open source Camino toolkit [10] to fit the models. Each model is fitted 250 times, using the Levenberg-Marquardt algorithm with a perturbed starting point from initial estimates drawn from the DT, to extract the parameters that produce the minimum objective function. The fitting uses an offset-Gaussian noise model to construct the objective $LSE = \sum_{i=1}^{N} \frac{1}{\sigma^2}(\tilde{S}_i - \sqrt{S_i^2 + \sigma^2})^2$, where N is the number of measurements, $\tilde{S}_i$ is the i-th measured signal, S_i its prediction from the model; σ^2 is the signal variance, which we estimate *a priori* from the b=0 images (this corresponds to an SNR of around 20). This objective function accounts for bias introduced by the Rician noise inherent in the MRI data in a way that is more numerically stable and computationally efficient than a full Rician log-likelihood objective function.

For model selection, we use the $BIC = -2\log(L) + K\log(N)$ where L is the likelihood of model parameters given the data, N is the number of measurements and K is the number of free parameters. We evaluate the BIC for each fitted model and then rank all models from lowest BIC (best) to highest (worst).

Table 1. Various model parameters from different data sets of angular thresholds of 2°, 5° and 10°. The models are ordered top-down by the BIC score of 2° data set. Here, we also include the estimates (shown in bold) from the best model of a previous ranking of non-dispersive parametric models [9]. [Note: Zepp=Zeppelin; ZepT=Zeppelin with tortuosity;Tens=Tensor; St=Stick; Bing=Bingham; Wat=Watson].

Param.		2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10
Nr	**Models**	**BIC**			**Intra.1 Vol.Fr.**			**Extra. Vol.Fr.**			**CSF/Dot Vol.Fr.**			**Axial Diff.**			**Radial Diff.**			**Kappa**		
10	Zepp.Bing.CSF.	513	380	359	0.56	0.59	0.60	0.29	0.30	0.31	0.15	0.11	0.09	2.0	1.9	1.9	0.5	0.6	0.7	6.9	7.1	6.5
9	ZepT.Bing.CSF.	516	377	356	0.59	0.59	0.60	0.28	0.30	0.31	0.13	0.11	0.09	2.0	1.9	1.9	0.6	0.7	0.7	7.0	7.1	6.5
12	Tens.Bing.CSF.	516	383	362	0.56	0.59	0.60	0.29	0.30	0.31	0.15	0.11	0.09	2.0	1.9	1.9	0.6	0.7	0.8	8.1	8.0	7.3
10	Tens.Wat.CSF.	519	392	369	0.55	0.59	0.60	0.29	0.29	0.30	0.16	0.12	0.10	2.0	1.9	1.9	0.5	0.7	0.7	5.4	5.5	5.3
8	Zepp.Wat.CSF.	531	401	373	0.56	0.59	0.60	0.29	0.30	0.31	0.15	0.11	0.09	2.0	1.9	1.9	0.5	0.6	0.7	5.6	5.6	5.4
7	ZepT.Wat.CSF.	533	398	369	0.59	0.60	0.60	0.28	0.30	0.31	0.13	0.11	0.09	2.0	1.9	1.9	0.6	0.6	0.7	5.8	5.6	5.4
9	ZepT.Bing.Dot	542	367	342	0.53	0.52	0.52	0.43	0.44	0.44	0.03	0.04	0.04	2.1	2.0	2.0	0.9	0.9	0.9	9.4	10.4	9.3
10	Zepp.Bing.Dot	544	366	340	0.50	0.48	0.48	0.45	0.47	0.47	0.04	0.05	0.05	2.1	2.0	2.0	0.9	0.8	0.8	10.3	12.0	11.0
12	Tens.Bing.Dot	548	371	345	0.51	0.49	0.48	0.45	0.46	0.47	0.04	0.05	0.05	2.1	2.0	2.0	1.0	0.9	0.9	10.7	12.5	11.5
10	Tens.Wat.Dot	557	385	355	0.49	0.47	0.47	0.46	0.48	0.48	0.04	0.05	0.05	2.1	2.0	2.0	1.0	0.9	0.9	8.1	8.7	8.2
9	Zepp.Bing.	559	398	370	0.64	0.65	0.64	0.36	0.35	0.36				2.2	2.1	2.0	1.0	1.0	1.0	6.6	6.7	6.2
7	ZepT.Wat.Dot	559	390	357	0.53	0.52	0.52	0.44	0.44	0.44	0.03	0.04	0.04	2.1	2.0	2.0	1.0	0.9	0.9	7.4	7.7	7.2
8	Zepp.Wat.Dot	561	389	356	0.50	0.48	0.48	0.46	0.47	0.47	0.04	0.05	0.05	2.1	2.0	2.0	0.9	0.8	0.8	8.0	8.6	8.2
11	Tens.Bing.	561	399	372	0.65	0.65	0.65	0.35	0.35	0.35				2.2	2.1	2.1	1.2	1.1	1.1	7.2	7.4	6.8
9	Tens.Wat.	575	418	384	0.64	0.65	0.65	0.36	0.35	0.35				2.2	2.1	2.1	1.1	1.1	1.0	5.5	5.5	5.3
8	ZepT.Bing.	576	416	385	0.62	0.62	0.62	0.38	0.38	0.38				2.2	2.1	2.1	0.8	0.8	0.8	6.5	6.5	6.0
7	Zepp.Wat.	576	419	383	0.64	0.65	0.65	0.36	0.35	0.35				2.2	2.1	2.1	1.0	1.0	1.0	5.6	5.5	5.3
6	ZepT.Wat.	593	437	398	0.62	0.63	0.63	0.38	0.37	0.37				2.2	2.1	2.1	0.8	0.8	0.8	5.4	5.3	5.1
12	Tens.St.St.Dot	652	464	439	0.23	0.22	0.21	0.56	0.56	0.56	0.07	0.07	0.08	2.0	1.9	1.9	0.8	0.8	0.8			
10	Zepp.St.St.Dot	658	464	437	0.23	0.22	0.21	0.56	0.56	0.56	0.07	0.07	0.08	2.0	1.9	1.9	0.8	0.7	0.7			
12	Tens.St.St.CSF.	674	562	557	0.22	0.24	0.23	0.41	0.41	0.42	0.21	0.18	0.17	1.5	1.5	1.4	0.5	0.5	0.5			
10	Zepp.St.St.CSF.	692	570	565	0.23	0.25	0.24	0.41	0.41	0.42	0.21	0.18	0.17	1.5	1.4	1.4	0.4	0.4	0.5			
8	Ball.Bing.	729	590	583	0.72	0.71	0.71	0.28	0.29	0.29				2.2	2.1	2.1				6.0	6.2	5.9
9	Ball.Bing.Dot	732	593	586	0.72	0.71	0.71	0.28	0.29	0.29	0.00	0.00	0.00	2.2	2.1	2.1				6.0	6.2	5.9
9	Ball.Bing.CSF.	732	593	586	0.72	0.71	0.71	0.28	0.29	0.29	0.00	0.00	0.00	2.2	2.1	2.1				6.0	6.2	5.9
6	Ball.Wat.	745	610	596	0.72	0.72	0.71	0.28	0.28	0.29				2.2	2.1	2.1				5.2	5.3	5.2
7	Ball.Wat.CSF.	748	613	599	0.72	0.72	0.71	0.28	0.28	0.29	0.00	0.00	0.00	2.2	2.1	2.1				5.2	5.3	5.2
7	Ball.Wat.Dot	748	613	599	0.72	0.72	0.71	0.28	0.28	0.29	0.00	0.00	0.00	2.2	2.1	2.1				5.2	5.3	5.2
7	**Zepp.St.Dot**	**784**	**597**	**570**	**0.29**	**0.30**	**0.29**	**0.62**	**0.62**	**0.62**	**0.09**	**0.09**	**0.09**	**1.9**	**1.9**	**1.8**	**0.7**	**0.7**	**0.7**			
10	**Tens.Cyl.CSF.**	**832**	**735**	**739**	**0.29**	**0.31**	**0.31**	**0.47**	**0.47**	**0.48**	**0.24**	**0.22**	**0.21**	**1.3**	**1.3**	**1.3**	**0.3**	**0.4**	**0.4**			
9	ZepT.St.St.Dot	843	652	640	0.33	0.33	0.32	0.50	0.50	0.50	0.05	0.05	0.05	1.8	1.7	1.7	1.1	1.1	1.1			
11	Tens.St.St.	859	687	666	0.28	0.28	0.27	0.52	0.51	0.52				1.7	1.6	1.6	0.9	0.9	0.9			
9	Zepp.St.St.	874	695	674	0.29	0.28	0.27	0.52	0.51	0.52				1.6	1.6	1.6	0.8	0.8	0.8			
9	ZepT.St.St.CSF.	881	718	712	0.35	0.35	0.34	0.39	0.41	0.43	0.12	0.10	0.08	1.5	1.5	1.4	0.9	0.9	0.9			
8	ZepT.St.St.	895	713	696	0.32	0.31	0.30	0.50	0.49	0.50				1.6	1.6	1.5	1.0	1.0	1.0			
8	Ball.St.St.	1,161	978	965	0.28	0.27	0.27	0.46	0.46	0.46				1.5	1.5	1.5						
9	Ball.St.St.Dot	1,162	977	965	0.24	0.26	0.25	0.47	0.46	0.47	0.02	0.02	0.02	1.6	1.5	1.5						
9	Ball.St.St.CSF.	1,164	981	969	0.28	0.27	0.26	0.46	0.46	0.46	0.00	0.00	0.00	1.5	1.5	1.5						

3 Results

Table 1 lists main parameter estimates across all three datasets (with η=2°, 5° and 10°). We have also included the best model from a previous ranking of non-dispersive parametric models [9], and a similar model with CSF instead of Dot. Four groups can be distinguished:

i) all combinations that include an anisotropic extracellular compartment and a Bingham/Watson intracellular compartment;

ii) models similar to (i) but instead using two-Sticks for their intracellular compartment, excluding models that use tortuosity or those without a spherically restricted compartment;

iii) all models incorporating an isotropic extracellular compartment with a Bingham/Watson intracellular compartment; and

iv) all exceptions to two-Sticks models in (ii).

The models that include a Bingham/Watson distribution outperform two-Sticks ones not simply because of their good quality of fit to the data but also reduced complexity.

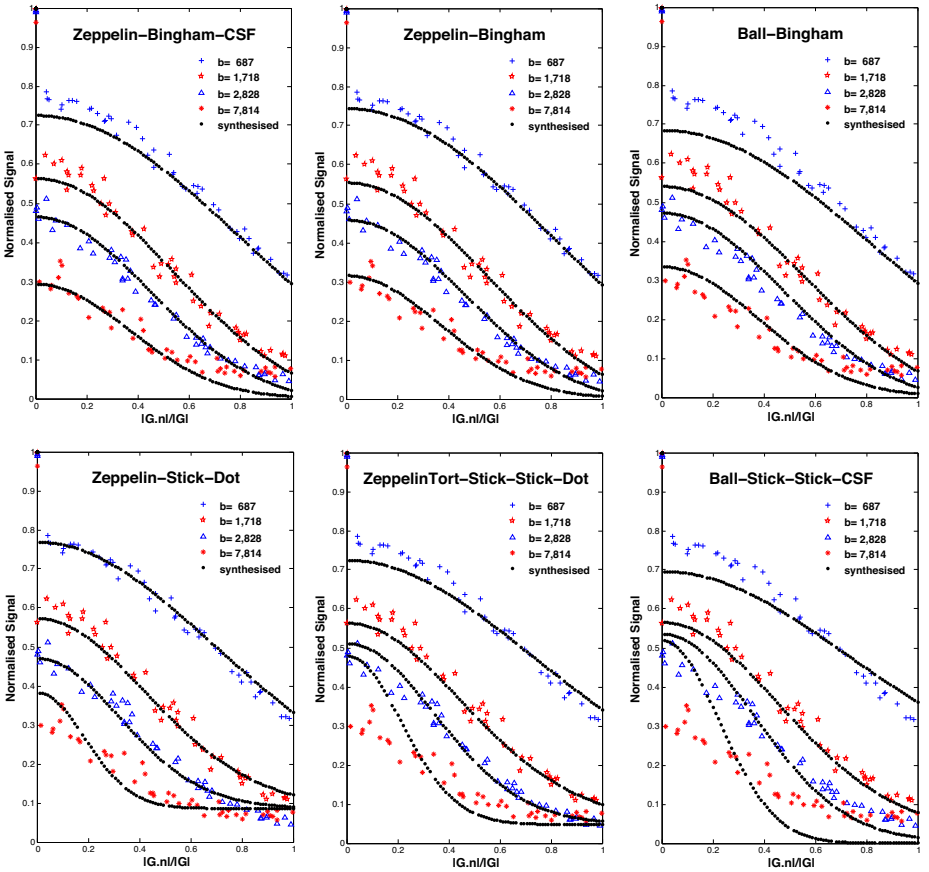

Fig. 2. A comparison of raw vs. predicted/synthesised signal from six representative models. The models are ordered in decreasing ranking left-right, top-bottom.

Within group (i), CSF models perform best for η=2° but, as η increases, Dot models are best. In this group, models using tortuosity produce similar estimates to those of the unconstrained Zeppelin, suggesting that meaningful constraints on the model parameters simplify the problem at little cost to fitting quality.

Across angular thresholds, the axial diffusivity is about $2\text{x}10^{-9}$ mm^2/s, and the radial diffusivity is around one-quarter of this in models with CSF, but one-half in others; this is to be expected as the CSF compartment has a fixed diffusivity of $3\text{x}10^{-9}$ mm^2/s and higher volume fraction than Dot.

As η increases from 2° to 5°, all models reflect the signal improvement from averaging across more voxels (24 vs. 66, resp.) through decreasing BIC and increasing fibre incoherence κ; however, at 10° (with 99 voxels averaged), the fitting improves slightly, but κ reflects the increased fibre coherence through decreasing κ.

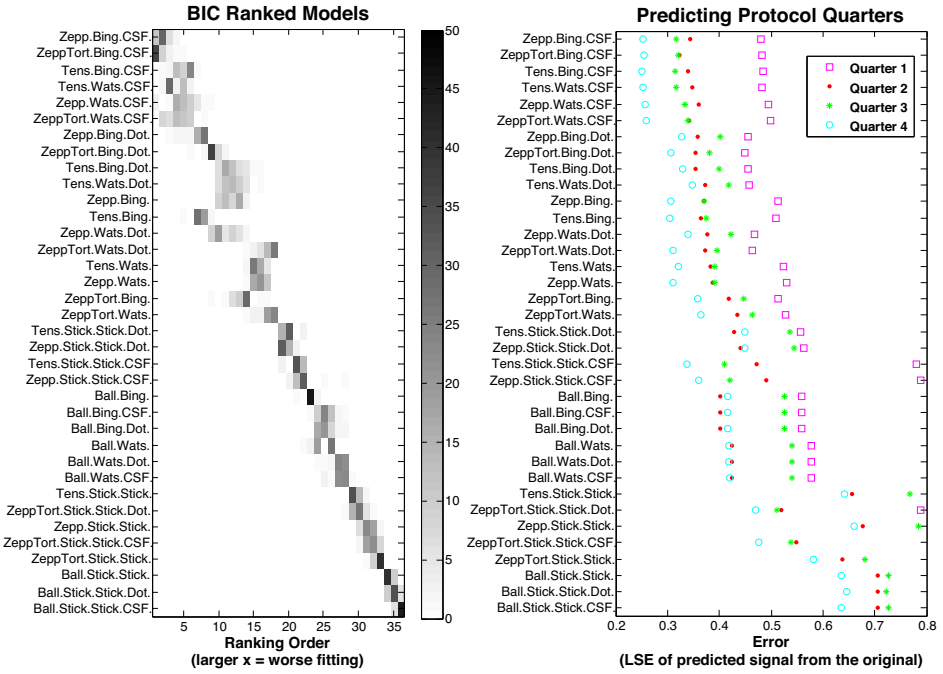

Fig. 3. LEFT: Positional variance diagrams over 100 bootstraps from the 2° data sets. The frequency of x-axis ranking is given by the shade of grey.
RIGHT: The accuracy of predicting unseen quarters of the protocol using parameters fitted to data from the remaining three-quarters. The ranking is as in Table 1.

Fig.2 shows the fit of some representative models to the data, to illustrate the difference between the actual signal and that generated from the model.

Fig.3 shows on the left the positional variance diagram for the *BIC* ranking through classical bootstrap. Each bootstrap data set is constructed through a random selection in each shell of the same number of data points, with replacement. The group structure remains unchanged, though there are minor variations within each group.

On the right of Fig.3, the relative performance of each model in reproducing unseen parts of the dataset is shown. We divide the data into four quarters, by randomly assigning low and high Δs into four groups. Then, we choose signal coming from three-quarters of the dataset to fit our models to and, from the parameter estimates drawn from these quarters, synthesise signal for the missing part. Next, we evaluate the sum of squared differences *LSE* compared to that unseen quarter. This provides an alternative model selection routine, to confirm and validate the ranking by *BIC* and, broadly speaking, the trends in both techniques agree.

4 Discussion

In this preliminary work, we have shown the advantage of dispersion models in describing data even in a homogeneous region of the brain such the CC.
In such structure, where a multitude of function specific fibre tracts bundle together, there is inhomogeneity that can produce a dispersion pattern, which is something that these models may reflect. In particular, the single mode orientation distributions (Watson/Bingham) outperform two discrete orientations (two-Sticks). As in previous work, an anisotropic extracellular compartment benefits the fitting, as does the addition of an isotropically restricted compartment.

We acknowledge that averaging voxels across parts of the CC and minor misalignments during image registration may exaggerate the dispersion, so this will be our future work. Testing the models voxel-by-voxel will help us see how consistent the model selection is within regions. We also intend to extend the investigation to other white matter structures that have greater dispersion.

Acknowledgements. EPSRC support this work through grants EP/E007748 and EP/I027084/01. We thank the MS Society of Great Britain for supporting the NMR unit. This work was supported by National Institute of Health Research and UCLH.

References

1. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson. B. 103, 247–254 (1994)
2. Stanisz, G.J., Szafer, A., Wright, G.A., Henkelman, M.: An analytical model of restricted diffusion in bovine optic nerve. Magn. Reson. Med. 37, 103–111 (1997)
3. Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, J.M.: Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088 (2003)
4. Zhang, H., Schneider, T., Wheeler-Kingshott, C.A.M., Alexander, D.C.: Practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage 61, 1000–1016 (2012)
5. Sotiropoulos, S., Behrens, T.E., Jbabdi, S.: Ball and Rackets: Inferring Fibre Fanning from Diffusion-weighted MRI. NeuroImage 60, 1412–1425 (2012)
6. Panagiotaki, E., Schneider, T., Siow, B., Hall, M.G., Lythgoe, M.F., Alexander, D.C.: Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. Neuroimage 59, 2241–2254 (2012)
7. Szafer, A., Zhong, J.H., Gore, J.C.: Theoretical model for water diffusion in tissues. Magn. Reson. Med. 33, 697–712 (1995)
8. Mardia, K.V., Jupp, P.E.: Distributions on spheres. Directional Stats., 159–192 (2000)
9. Ferizi, U., Schneider, T., Panagiotaki, E., Nedjati-Gilani, G., Zhang, H., Wheeler-Kingshott, C.A.M., Alexander, D.C.: Ranking Diffusion MRI Models with In Vivo Human Brain Data. In: 10th IEEE International Symposium on Biomedical Imaging (2013)
10. Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J., Alexander, D.C.: Camino: Open-source diffusion-MRI reconstruction and processing. In: 14th Scientific Meeting of the ISMRM 2759 (2006)