Maximum Entropy Spherical Deconvolution for Diffusion MRI

Daniel C Alexander[1]

Department of Computer Science.
University College London, Gower Street, London, WC1E 6BT, UK.
D.Alexander@cs.ucl.ac.uk

Abstract. This paper proposes a maximum entropy method for spherical deconvolution. Spherical deconvolution arises in various inverse problems. This paper uses the method to reconstruct the distribution of microstructural fibre orientations from diffusion MRI measurements. Analysis shows that the PASMRI algorithm, one of the most accurate diffusion MRI reconstruction algorithms in the literature, is a special case of the maximum entropy spherical deconvolution. Experiments compare the new method to linear spherical deconvolution, used previously in diffusion MRI, and to the PASMRI algorithm. The new method compares favourably both in simulation and on standard brain-scan data.

1 Introduction

Diffusion MRI has exploded over the last decade, since the introduction of diffusion-tensor MRI (DT-MRI) by Basser et al [1]. Diffusion MRI measures the displacement of particles, usually water molecules, within a material over a fixed time interval. The material microstructure controls the scatter pattern of particles within and, conversely, measurements of the particle displacement reveal information about the microstructure. The current standard diffusion MRI technique is DT-MRI, which provides two unique insights into material microstructure. First, DT-MRI provides quantitative measurements of the anisotropy of particle displacements and, second, it provides an estimate of the dominant orientation of particle displacements. In fibrous material, such as white matter in the brain, the dominant orientations of particle displacements are similar to the dominant fibre directions. Diffusion MRI is particularly useful for brain imaging, because it reveals the orientations of white-matter fibres in each voxel of an image volume. Tractography algorithms use these fibre-orientation estimates to determine the connectivity of the whole brain.

A well-documented problem with DT-MRI is that it fails at fibre crossings. Recent trends in the field are towards a new generation of reconstruction algorithm that can resolve the orientations of multiple fibre populations within each voxel, see [2] for a review. One such technique [3, 4] views the diffusion MRI signal as the convolution of the response from a single fibre with specific orientation with the distribution of fibre orientations. A deconvolution of the signal

by the single-fibre response yields the fibre-orientation distribution (FOD). The implementations of this deconvolution technique in [3, 4] use a linear basis for spherical functions to represent the FOD. Although the literature shows nice results from the technique in certain regions of high-quality data, the methods are reknowned for instability and, in performance comparisons such as [5], they perform worse than rival methods such as PASMRI [6] and **q**-ball imaging [7].

Maximum entropy methods [8] have proved useful in a variety of reconstruction and inverse problems. In particular, the methods have proved effective for deconvolution. This paper constructs a maximum-entropy formulation of the spherical deconvolution problem and demonstrates its application to reconstruction of the FOD from diffusion MRI measurements. Experiments show that the maximum-entropy spherical-deconvolution (MESD) improves on a linear implementation and produces results comparable to the PASMRI algorithm on data from a standard diffusion MRI acquisition sequence. Further analysis shows that the PASMRI algorithm is a special case of the MESD method. Performance of the method depends on the choice of response function (deconvolution kernel). Here we use only a simple kernel to show efficacy of the approach. We can expect further improvements in the method through better choices of deconvolution kernel. Although particularly useful for diffusion MRI, the method extends naturally to any spherical deconvolution problem.

Section 2 gives some background on the diffusion MRI reconstruction problem and existing techniques. Section 3 outlines the MESD method and compares it analytically to other reconstruction techniques used in diffusion MRI. Section 4 shows results from using the new methods on diffusion MRI data acquired from a standard brain imaging sequence and compares them with results from linear spherical deconvolution (SD) and PASMRI on the same data. Simulation experiments compare the methods further. Finally, section 5 concludes and outlines areas for further work.

2 Background

Diffusion-tensor MRI models the particle displacements in each image voxel with a zero-mean Gaussian distribution with covariance $2t\mathsf{D}$, where t is the diffusion time and D is the diffusion tensor. The diffusion MRI measurement is approximately proportional to the Fourier transform of the particle-displacement density function p:

$$A(\mathbf{q}) = (A^\star(\mathbf{0}))^{-1} A^\star(\mathbf{q}) = \int_{\mathbb{R}^3} p(\mathbf{x}) \cos(\mathbf{q} \cdot \mathbf{x}) \mathrm{d}\mathbf{x}, \tag{1}$$

where $A^\star(\mathbf{q})$ is the MR signal with wavenumber $\mathbf{q}$, which depends on the strength and direction of the magnetic gradient pulses in the diffusion MR imaging sequence [9]. The MRI measurements also vary spatially, although we omit the spatial dependence from the notation for simplicity. Substituting the Gaussian model for p in (1) and taking logs shows that each measurement provides a linear constraint on the six elements of the diffusion tensor, D. We can estimate D from seven or more measurements with independent $\mathbf{q}$. Since the contours of

the Gaussian model are ellipsoidal, DT-MRI can only reveal a single dominant fibre orientation in each voxel. At fibre crossings, however, p has multiple peaks at fixed radii and the Gaussian model is poor. We can also fit other models for p using (1), such as mixtures of Gaussians, e.g. [10,11], which can reveal the directions of multiple fibre orientations.

Although DT-MRI extends to handle multiple fibre populations through the use of mixture models, fitting becomes non-linear with mixtures of two or more Gaussians and is often unstable. Furthermore, model fitting procedures require a choice of the number of fibres in each voxel prior to fitting. A new generation of *multiple-fibre reconstruction* techniques [2,12], including PASMRI [6], **q**-ball imaging [7] and spherical-deconvolution methods [3,4] all compute objects that provide the number of fibres, together with an orientation estimate for each, without requiring prior specification of the number of fibres present.

Spherical-deconvolution methods assume that the diffusion MRI signal in each voxel is the convolution of the FOD (fibre-orientation distribution) f, which is a real-valued function of the unit sphere, with the signal $R(\cdot;\hat{\mathbf{x}})$ from a single fibre with orientation $\hat{\mathbf{x}}$:

$$A(\mathbf{q}) = \int R(\mathbf{q};\hat{\mathbf{x}}) f(\hat{\mathbf{x}}) \mathrm{d}\hat{\mathbf{x}}, \tag{2}$$

where the integration is over the unit sphere in three-dimensional space. Note that (2) assumes that $R(\cdot;\hat{\mathbf{x}})$ has rotational symmetry about $\hat{\mathbf{x}}$ and does not vary spatially. The methods aim to deconvolve the signal, using a model for $R(\cdot;\hat{\mathbf{x}})$, to obtain f. The function f can have multiple pairs of equal and opposite peaks and each pair provides a separate fibre-orientation estimate.

To implement the method, the standard approach represents f using a linear basis:

$$f(\hat{\mathbf{x}}) = \sum_{k=1}^{K} \beta_k \theta_k(\hat{\mathbf{x}}). \tag{3}$$

We substitute for f in (2) and reverse the order of the integral and sum to obtain

$$A(\mathbf{q}) = \sum_{k=1}^{K} \left(\beta_k \int R(\mathbf{q};\hat{\mathbf{x}}) \theta_k(\hat{\mathbf{x}}) \mathrm{d}\hat{\mathbf{x}} \right). \tag{4}$$

Diffusion MRI sequences usually acquire a set of measurements with wavenumbers $\mathbf{q}_i$, $i = 1, \cdots, N$, together with some number of measurements with $\mathbf{q} = \mathbf{0}$ for normalization. The set of $\mathbf{q}_i$ is the same in each voxel. For a set of measurements with wavenumbers $\mathbf{q}_i$, $i = 1, \cdots, N$, we can summarize the set of equations from (4) as $\mathbf{A} = \mathbf{XB}$, where $\mathbf{A} = (A(\mathbf{q}_1), \cdots, A(\mathbf{q}_N))^T$ is the vector of normalized measurements, $\mathbf{B} = (\beta_1, \cdots, \beta_K)^T$ is the vector of basis-function weights and $\mathbf{X}$ is the matrix with ik-th entry

$$X_{ik} = \int R(\mathbf{q}_i;\hat{\mathbf{x}}) \theta_k(\hat{\mathbf{x}}) \mathrm{d}\hat{\mathbf{x}}.$$

We solve the matrix equation to obtain the set of basis-function weights that define f via a linear transformation of the measurements: $\mathbf{B} = \mathbf{X}'\mathbf{A}$, where $\mathbf{X}' = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is the pseudoinverse of $\mathbf{X}$. Since the set of $\mathbf{q}_i$ is identical in each voxel, we need to compute $\mathbf{X}'$ only once. The computational burden of the method is therefore light (a single matrix multiplication in each voxel) and comparable to that of diffusion-tensor MRI.

Following the general method outlined in [13], references [3, 4] use the spherical harmonics as the basis for f. References [3, 11] use Gaussian models of the particle displacement within single fibres to obtain R. Tournier et al [4] derive R directly from the data by taking an average signal from the most anisotropic voxels.

The PASMRI algorithm [6] aims to compute a feature of p called the persistent angular structure (PAS) by assuming independence of the angular and radial structure of p, so that $p(\mathbf{x}) = g(|\mathbf{x}|)\hat{p}(\hat{\mathbf{x}})$, where g is a model for the radial structure of p and $\hat{p}$ is the PAS. Jansons and Alexander [6] take $g(|\mathbf{x}|) = \delta(|\mathbf{x}|-r)$ for some scalar r, so that (1) becomes

$$A(\mathbf{q}_i) = r^{-2} \int \hat{p}(\hat{\mathbf{x}}) \cos(r\mathbf{q}_i \cdot \hat{\mathbf{x}}) \mathrm{d}\hat{\mathbf{x}}. \tag{5}$$

Jansons and Alexander derive a maximum-entropy parametrization of $\hat{p}$:

$$\hat{p}(\hat{\mathbf{x}}) = \exp\left(\lambda_0 + \sum_{j=1}^{N} \lambda_j \cos(r\mathbf{q}_j \cdot \hat{\mathbf{x}})\right), \tag{6}$$

which they fit to the measurements in each voxel using a Levenberg–Marquardt algorithm. The recovered PAS reflects the angular structure of p that persists over a wide range of radii. Like the FOD, the PAS is a real-valued function of the sphere with peaks that provide fibre-orientation estimates.

Another class of method, including **q**-ball imaging [7], diffusion spectrum imaging [14], and the methods in [15, 16], all compute or estimate the *orientation distribution function* (ODF) ϕ, which is the projection of p onto the unit sphere:

$$\phi(\hat{\mathbf{x}}) = \int_0^\infty p(\alpha\hat{\mathbf{x}}) \mathrm{d}\alpha. \tag{7}$$

The ODF is also a real-valued function of the sphere with peaks that provide fibre-orientation estimates. If we represent ϕ using a linear basis, recovery of the ODF in each voxel comes from multiplication of the measurements by the same matrix in each voxel so computation time is similar to linear SD, see [7, 12].

3 Maximum Entropy Spherical Deconvolution

This section outlines the new MESD method. Since the literature already contains a variety of multiple-fibre reconstruction algroithms, we begin with some motivation for the development of the new technique.

The literature contains little evaluation and comparison of multiple-fibre reconstructions, but early indications [5, 12] suggest that the PASMRI algorithm recovers fibre directions more accurately and consistently than spherical deconvolution and the **q**-ball algorithm. The PASMRI algorithm has two fundamental differences to the other methods that may account for differences in performance: it computes a different object, the PAS, and it uses a non-linear basis to represent that object. Comparison of Eq. (2) with Eq. (5) reveals that the PASMRI inversion is a deconvolution with $R(\mathbf{q}; \hat{\mathbf{x}}) = r^{-2}\cos(r\mathbf{q} \cdot \hat{\mathbf{x}})$. If we replace the maximum-entropy representation of the PAS in Eq. (6) with a linear combination of basis functions, as used for f in Eq. (3), the method reduces to a single matrix multiplication in each voxel like the linear SD method. However, experiments with this linearized version of PASMRI (not shown) reveal that performance is signficantly worse than the non-linear implementation in [6]. This suggests that the power of the method lies in the non-linear maximum-entropy representation of the PAS. In fibrous tissue, the PAS function often has very sharp peaks, which linear bases cannot capture accurately but the product of exponential waves in Eq. (6) is better equipt to represent.

Spherical-deconvolution methods have some advantages over the PASMRI method and methods that estimate the ODF. First, the output of the spherical deconvolution, the FOD, is a readily understandable object with well-defined meaning. The PAS is more arcane and it is less obvious why its peaks should correspond to fibre orientations. Second, spherical deconvolution does not rely on the Fourier relationship between the MRI measurements and p, which is only approximate. Third, the peaks of the FOD, at least in theory using an ideal deconvolution kernel, correspond genuinely to fibre orientations rather than the directions of ridges in p. The peaks of p at a fixed radius may not correspond exactly to fibre directions particularly if the fibre orientations are not orthogonal. If p is a mixture of Gaussians, for example, the peaks at fixed radii (or ridge directions) are more closely aligned than the peaks of the individual Gaussian components when the peaks are not orthogonal. The effect is similar in mixtures of Gaussians in one dimension where the peaks of the mixture are closer than those of the mixed components.

The observation that the linear representation of the PAS produces worse performance than the non-linear representation suggests that the poor performance of spherical-deconvolution methods in diffusion MRI may arise from the linear basis representation of the FOD. Like the PAS, we may expect the FOD to have sharp peaks in fibrous tissue regions where the distribution of fibre directions is highly concentrated. This section derives an alternative non-linear representation for the FOD using a maximum-entropy argument similar to that used by Jansons and Alexander [6] to derive the PAS representation in Eq. (6).

We aim to determine a representation for f that imposes the minimum information on the reconstructed f. The information content of the FOD, f, is

$$I[f] = \int f(\hat{\mathbf{x}}) \log f(\hat{\mathbf{x}}) \, \mathrm{d}\hat{\mathbf{x}}. \tag{8}$$

We shall minimize the information content of f subject to the constraints from the measurements and that f is a probability density function and so integrates to one:

$$\int f(\hat{\mathbf{x}})\mathrm{d}\hat{\mathbf{x}} = 1. \tag{9}$$

Each measurement provides a constraint on f given by Eq. (2). We incorporate each constraint into the expression for the information content of f using the method of Lagrange multipliers to yield

$$I[f] = \int \left(f(\hat{\mathbf{x}}) \log f(\hat{\mathbf{x}}) - f(\hat{\mathbf{x}}) \sum_{i=1}^{N} (\lambda_i R(\mathbf{q}_i; \hat{\mathbf{x}})) - f(\hat{\mathbf{x}})\mu \right) \mathrm{d}\hat{\mathbf{x}}, \tag{10}$$

where $\mathbf{q}_i$, $i = 1, \cdots, N$, are the wavenumbers of the MRI measurements, the λ_i are Lagrange multipliers for the constraints from the data and the Lagrange multiplier μ controls the normalization of f. Taking a variational derivative $\delta I[f]$ and solving $\delta I[f] = 0$, we find that the information content, $I[f]$, is minimum when

$$f(\hat{\mathbf{x}}) = \exp\left(\lambda_0 + \sum_{i=1}^{N} \lambda_i R(\mathbf{q}_i; \hat{\mathbf{x}}) \right), \tag{11}$$

where $\lambda_0 = \mu - 1$.

We need to solve

$$\int f(\hat{\mathbf{x}}) R(\mathbf{q}_i; \hat{\mathbf{x}}) \mathrm{d}\hat{\mathbf{x}} = A(\mathbf{q}_i) \tag{12}$$

for the λ_i. We implement the method following Jansons and Alexander's implementation of the PASMRI algorithm in [6]. We use a Levenburg–Marquardt algorithm to search for a set of λ_i that minimize

$$\sum_{i=1}^{N} \left(A(\mathbf{q}_i) - \int f(\hat{\mathbf{x}}) R(\mathbf{q}_i; \hat{\mathbf{x}}) \mathrm{d}\hat{\mathbf{x}} \right)^2. \tag{13}$$

In all the experiments in the next section we use the simple deconvolution kernel

$$R(\mathbf{q}; \hat{\mathbf{x}}) = \exp\left(-t|\mathbf{q}|^2 d^{-1} (\hat{\mathbf{x}} \cdot \hat{\mathbf{q}})^2 \right), \tag{14}$$

following [11], where d is the diffusivity in the fibre direction; we take $t|\mathbf{q}|^2 d^{-1} = 1$, since the $|\mathbf{q}_i|$ are all equal in the test data we use. The kernel in Eq. (14) is the signal we expect from a material in which particles displace only in direction $\hat{\mathbf{x}}$.

4 Experiments and Results

We begin by comparing the output of the MESD algorithm with linear SD and the PASMRI algorithm. Figure 1(a) shows the maximum entropy FOD in each voxel of the image region spanning the two highlighted regions of interest on the

coronal slice in figure 2. Figure 1(b) shows the linear FOD over the same region and figure 1(c) shows the PAS. The diffusion MRI data comes from a standard acquisition sequence with $N = 54$, $t = 0.04\,\mathrm{s}$, $|\mathbf{q}_i| = 2.00 \times 10^5\,\mathrm{m}^{-1}$, $i = 1, \cdots, N$, and the $\hat{\mathbf{q}}_i$ minimize the electrostatic energy with equal charges at each $\hat{\mathbf{q}}_i$ and $-\hat{\mathbf{q}}_i$. The signal to noise ratio with $|\mathbf{q}| = \mathbf{0}$ in white matter is approximately 16, which is lower than for the test data used to introduce many multiple-fibre reconstructions in the literature but is typical for whole-brain diffusion MRI acquisitions.

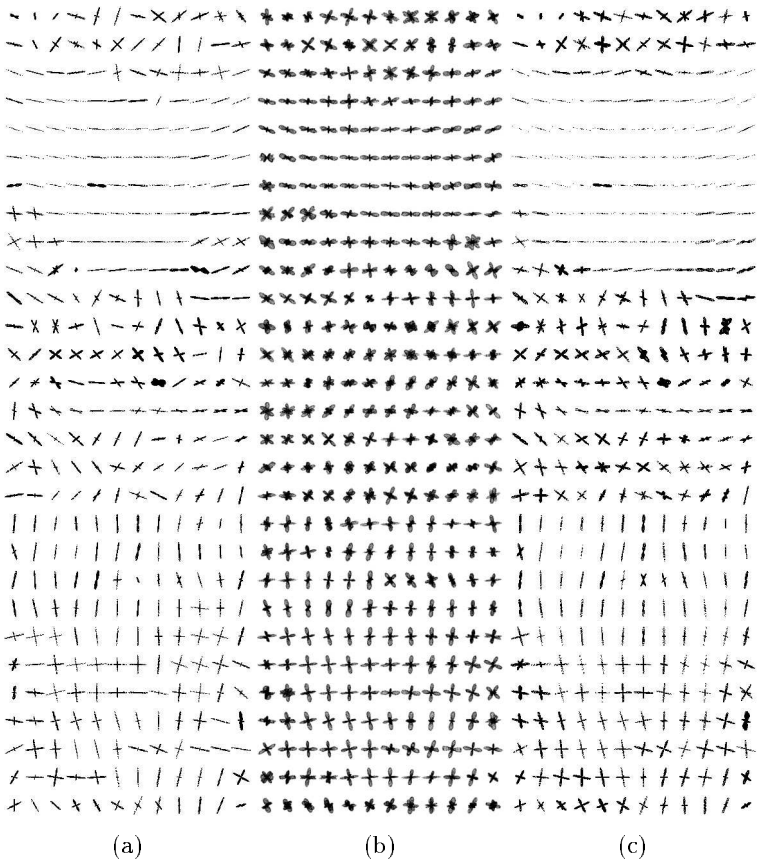

Fig. 1. Reconstructions over the region spanning the two highlighted regions in figure 2 for (a) MESD and (b) linear SD and (c) PASMRI.

In the linear SD, we use a radial-basis-function representation with

$$\theta_k(\hat{\mathbf{x}}) = \exp(-(\cos^{-1}(|\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}_k|))^2/\sigma^2), \tag{15}$$

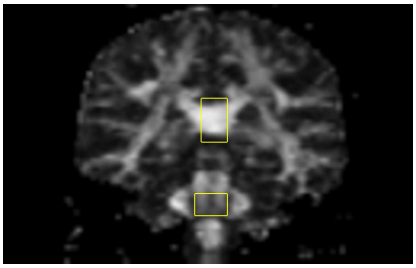

Fig. 2. Fractional anisotropy over a coronal slice of the brain data set with two highlighted regions of interest: the corpus callosum (top) and the fibre crossing at the pons (bottom).

where σ is a constant scaling parameter and the $\hat{\mathbf{y}}_k$, $k = 1, \cdots, K$, are unit vectors evenly distributed on the unit sphere. The method is simpler to control using radial basis functions than spherical harmonics. We take $K = 120$ and $\sigma = 20.0$. The value of σ controls the smoothness of the output of the linear SD method. With $\sigma > 50$, the basis functions are too broad for the linear combination to support two orthogonal peaks reliably. With $\sigma < 10$ spurious peaks dominate the real peaks both in simulation and on the brain data and the method gives no sensible output. In the PASMRI algorithm, we set $r = 1.4$ following the suggestion in [6].

The image region in figure 1 contains part of the corpus callosum (top), where we expect a single fibre with left-right orientation, and part of the pons, where we expect two approximately orthogonal fibres with left-right and superior-inferior orientations. All three methods produce peaks in the expected fibre directions in both regions. In the corpus callosum, the linear SD method consistently produces false-positive fibre directions approximately orthogonal to the correct fibre direction; PASMRI and MESD rarely generate false positives in the corpus callosum. In the pons, PASMRI and linear SD show two peaks more consistently than MESD. However, linear SD produces extra peaks pointing in the posterior-anterior direction (out of the page) in a significant proportion of voxels in the pons region; PASMRI produces similar false positives to a lesser extent; MESD rarely shows these false positives.

Further experiments compare the methods in simulation. We synthesize measurements from the imaging sequence for the brain data using a simple test function for p, which models the particle displacement in tissue containing two distinct fibre orientations. We set

$$p(\mathbf{x}) = aG(\mathbf{x}; \mathcal{D}_1, t) + (1 - a)G(\mathbf{x}; \mathcal{D}_2, t) \tag{16}$$

where $a \in [0, 1]$ is the mixing parameter, $G(\cdot; \mathcal{D}, t)$ is a zero-mean trivariate Gaussian function with covariance matrix $2t\mathcal{D}$ and the diffusion tensors are

$$\mathcal{D}_1 = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_2) \text{ and} \tag{17}$$

$$\mathcal{D}_2 = \mathbf{R}^T(\phi)\mathrm{diag}(\lambda_2, \lambda_1, \lambda_2)\mathbf{R}(\phi), \tag{18}$$

where $\mathbf{R}(\phi)$ is a rotation about the z-axis through angle ϕ and $\mathrm{Tr}(\mathcal{D}_i) = \lambda_1 + 2\lambda_2 = 2.1 \times 10^{-9}\,\mathrm{m^2 s^{-1}}$ (typical for brain tissue).

For various settings of a, ϕ and λ_1, we synthesize 256 voxels of data with independent noise and estimate the fibre orientations using MESD, linear spherical deconvolution and PASMRI. We use the peak-finding algorithm in [6] to determine the fibre directions from each method in each trial. In each set of trials, we compute the following performance statistics:

- The consistency fraction C, which is the fraction of trials in which the reconstruction algorithm finds the right number of fibres (two) to within a small angular tolerance of the principal directions of $\mathcal{D}_1$ and $\mathcal{D}_2$. We set the tolerance to $\cos^{-1} 0.95$.
- The direction concentration of each population of recovered fibre directions. The concentration of a collection of directions $\hat{\mathbf{x}}_i$, $i = 1, ..., n$, is $\gamma(\kappa_1) = -\log(1-\kappa_1)$, where κ_1 is the largest eigenvalue of the mean dyadic tensor $\mathsf{Y} = n^{-1}\sum_{j=1}^{n} \hat{\mathbf{x}}_i\hat{\mathbf{x}}_i^T$. To group corresponding peak directions from separate trials, we cluster the directions to maximize the concentration of each population.
- The direction bias of each recovered population. The bias is $\mu_1 \cdot \mathbf{n}$, where μ_1 is the principal eigenvector of Y and $\mathbf{n}$ is the closest of the principal directions of $\mathcal{D}_1$ and $\mathcal{D}_2$.

Table 1 shows the performance statistics for each algorithm with each combination of settings in the test function. Some observations from table 1 are:

- The consistency fraction is much greater for PASMRI and MESD than for linear SD. The linear method produces false positives in most trials, although it rarely produces false negatives. Reducing σ can reduce the false positive rate slightly, but the effect is not significant until a sudden change when the basis functions become too smooth to support two peaks and the false negative count increases sharply.
- The peak directions are more consistent using MESD than linear SD, since the direction concentrations are higher for MESD than linear SD.
- The bias in the mean reconstructed directions is lower for MESD than linear SD apart from when the test function is very anisotropic (low λ_1). At very high anisotropy, some of the measurements hit the noise floor, which corrupts the reconstruction.
- The PASMRI method shows better results than both spherical deconvolutions. Both deconvolution methods show greater bias in the mean reconstructed directions than PASMRI, particular when the test-function principal directions are non-orthogonal. The consistency fraction is generally slightly greater for PASMRI than for MESD. Further examination of the output reveals that PASMRI tends to have a greater false-positive rate than MESD, which has greater false-negative rate. However, we note that the false-positive false-negative trade off is easy to control in PASMRI by varying the parameter r.

In fact, we might expect that the PASMRI algorithm performs better in this simulation. To generate the synthetic data we use the Fourier model in Eq. (1), which is also the basis of the PASMRI algorithm. Since both deconvolution algorithms show similar trends in the bias of the reconstructed directions, it seems likely that the bias is a product of the choice of deconvolution kernel. The deconvolution kernel we use in the deconvolution methods does not match the components of the test function. We might expect that the spherical-deconvolution methods are more robust to the departures from the Fourier model in the measurement process from which the brain data comes; see [9, 17] for some discussion of the nature of the departures from the Fourier model in diffusion MRI data.

$\lambda_1/10^{-12}$ $\mathrm{m^2\,s^{-1}}$	a	ϕ deg	MESD C	MESD $\gamma(\kappa_1)$	MESD α/deg	PAS C	PAS $\gamma(\kappa_1)$	PAS α/deg	Lin. SD C	Lin. SD $\gamma(\kappa_1)$	Lin. SD α/deg
100	0.5	0	1.000	5.9	0.2	1.000	5.6	0.2	0.012	3.9	1.7
100	0.5	22.5	0.980	5.5	9.2	0.758	2.5	10.4	0.008	4.0	7.9
100	0.6	0	1.000	5.6	0.2	1.000	5.3	0.2	0.024	3.7	1.9
100	0.6	22.5	0.964	5.2	12.6	0.832	3.4	9.0	0.020	3.6	11.5
300	0.5	0	0.996	4.7	0.4	0.996	4.3	0.4	0.063	3.0	3.2
300	0.5	22.5	0.773	4.0	10.0	0.949	3.7	1.9	0.020	2.5	10.0
300	0.6	0	0.996	4.4	0.4	1.000	4.2	0.5	0.043	2.7	4.1
300	0.6	22.5	0.727	3.8	13.2	0.906	3.7	3.6	0.020	2.3	13.4
500	0.5	0	0.504	2.3	3.1	0.270	2.1	2.2	0.012	1.7	8.6
500	0.5	22.5	0.176	2.0	15.1	0.211	1.8	9.8	0.008	1.5	17.4
500	0.6	0	0.492	2.3	2.5	0.273	2.0	2.1	0.012	1.6	9.0
500	0.6	22.5	0.156	1.9	15.2	0.180	1.6	12.8	0.012	1.5	21.0

Table 1. Summary of simulation results for the MESD, PASMRI and linear SD. The angle $\alpha = \cos^{-1}(\mu_1 \cdot \mathbf{n})$ is the bias in recovered directions. Each set of trials produces two direction concentrations, $\gamma(\kappa_1)$, and two bias angles, one from each component. The table shows the lowest direction concentration and the highest bias angle of each pair.

5 Conclusions

This paper introduces a maximum-entropy spherical-deconvolution method and demonstrates efficacy within the diffusion MRI application. The method improves on a linear implementation of the spherical deconvolution method. With the chosen deconvolution kernel, the method does not quite match the performance of PASMRI in simulation, but the spherical deconvolution method, which generalizes the PASMRI method, potentially has theoretic advantages, which warrant its use and further investigation. Many avenues of future work may extend and improve the basic MESD implementation here. Better choices of

deconvolution kernel surely exist. In [4], Tournier et al estimate R directly from the input data. In more recent work [18], they improve the robustness to noise of linear SD by choosing a deconvolution kernel that minimizes the entropy of the output FOD. The MESD method is simple to adapt to use the kernels in [4, 18] or any other kernel.

A disadvantage of MESD is that computation time is much greater than for linear SD and is similar to that of the PASMRI algorithm. The numerical integration in (12) dominates the computation time. The numerical integration scheme is naive, however, and approximates the integral with a summation over points evenly distributed over the sphere. A better approach would sample the integrand more densely around its sharp peaks and avoid unnecessary evaluations where the integrand is near zero.

Other possible areas for improvement are the optimization algorithms that fit the non-linear FOD to the measurements. We use a single starting point for the optimization and do not guarantee to find the global minimum of the least-squares objective function. Other choices for the objective function itself may also improve results. For example, we may extend the maximum-entropy analogy by maximizing the entropy of the FOD subject to the data constraints rather than minimizing the least-squares fit to the data. Other regularization techniques may also improve stability.

Acknowledgements

The author thanks Claudia Wheeler-Kingshott at the Institute of Neurology, UCL, for providing the human brain data used in this work.

References

1. Basser P J, Matiello J and Le Bihan D 1994 MR diffusion tensor spectroscopy and imaging *Biophysical Journal* **66** 259–67
2. Alexander D C 2005 An introduction to computational diffusion MRI: the diffusion tensor and beyond *Visualization and Image Processing of Tensor Fields* ed Weichert J and Hagen H (Springer)
3. Anderson A and Ding Z 2002 Sub-voxel measurement of fiber orientation using high angular resolution diffusion tensor imaging *Proc. 10th Annual Meeting of the ISMRM (Honolulu)* (Berkeley, USA: ISMRM) 440
4. Tournier J-D, Calamante F, Gadian D G and Connelly A 2004 Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution *NeuroImage* **23** 1176–1185
5. Alexander D C 2005 Monte-Carlo studies of multiple-fibre reconstructions for diffusion MRI *Proc. 13th Annual Meeting of the ISMRM (Miami)* (Berkeley, USA: ISMRM) Accepted
6. Jansons K M and Alexander D C 2003 Persistent Angular Structure: new insights from diffusion MRI data *Inverse Problems* **19** 1031–1046
7. Tuch D S 2004 Q-ball imaging *Magnetic Resonance in Medicine* **52** 1358–1372

8. Skilling J and Gull S F 1985 Algorithms and applications *Maximum Entropy and Bayesian methods in Inverse Problems* ed Smith C R and Grandy W T (Dordrecht: Reidel publishing company) 83–132
9. Callaghan P T 1991 *Principles of Magnetic Resonance Microscopy* (Oxford, UK: Oxford Science Publications)
10. Tuch D S 2002 *Diffusion MRI of Complex Tissue Structure* (Doctor of Philosophy in Biomedical Imaging at the Massachusetts Institute of Technology)
11. Behrens T E J, Woolrich M W, Jenkinson M, Johansen-Berg H, Nunes R G, Clare S, Matthews P M, Brady J M and Smith S M 2003 Characterization and propagation of uncertainty in diffusion-weighted MR imaging *Magnetic Resonance in Medicine* **50** 1077–1088
12. Alexander D C 2005 Multiple-fibre reconstruction algorithms for diffusion MRI *Annals of the NYAS* In Press
13. Healy D M, Hendriks H and Kim P T 1998 Spherical deconvolution *Journal of Multivariate Analysis* **67** 1–22
14. Wedeen V J, Reese T G, Tuch D S, Dou J-G, Weiskoff R M and Chessler D 1999 Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI *Proc. 7th Annual Meeting of the ISMRM (Philadelphia)* (Berkeley, USA: ISMRM) 321
15. Lin C P, Tseng W Y I, Kuo L, Wedeen V J and Chen J H 2003 Mapping orientation distribution function with spherical encoding *Proc. 11th Annual Meeting of the ISMRM (Toronto)* (Berkeley, USA: ISMRM) 2120
16. Ozarslan E, Vemuri B C and Mareci T H 2004 Fiber orientation mapping using generalized diffusion tensor imaging *Proc. IEEE International Symposium on Biomedical Imaging (Arlington)* (IEEE)
17. Mitra P P and Halperin B I 1995 Effects of finite gradient-pulse widths in pulsed-field-gradient diffusion measurements *Journal of Magnetic Resonance* **113** 94–101
18. Tournier J-D, et al 2005 *Proc. 13th Annual Meeting of the ISMRM (Miami)* (Berkeley, USA: ISMRM) Accepted