
White matter fibre tracking

THE LOCAL white matter orientation information provided indirectly by dMRI can be used to reconstruct the pathways of major white matter structures through the brain. This reconstruction process is known as fibre tracking, or tractography. A considerable number of tractography algorithms have been put forward, however, which differ in the ways that they interpret the original data, how they handle uncertainty, and how they represent the reconstructed tract. In some cases nontensorial models of diffusion have been employed to handle some of the degeneracies that the diffusion tensor model faces.

In this chapter we review a number of different types of tractography algorithm, describe their relative advantages and disadvantages, and discuss some of the uses to which fibre tracking methods have been applied. We also mention some of the limitations that still apply to the state of the art algorithms.

5.1 Streamlines

We have seen in chapter 4 that the tensor model of diffusion provides an indication of the principal orientation and magnitude of diffusion at a point, in the form of the first eigenvector and associated eigenvalue. This information can be visualised simply by drawing a line, whose orientation and length indicate these two properties, at each location where the model is evaluated—typically a voxel. The components of this representation in a single axial (x - y) plane are shown in Fig. 5.1(a). It can be seen by inspection from this figure that there is a fairly smooth curvature in successive principal diffusion direction vectors as they progress across, in this case, the corpus callosum splenium. The most intuitive way to reconstruct a tract is, then, to link these directions together to form a **streamline**. This is the approach taken, in some form, by a majority of tractography algorithms.

Fig. 5.1(b) demonstrates the tract reconstruction process of the Fibre Assignment by Continuous Tracking (FACT) algorithm, which was first demonstrated for fixed rat brain tissue (Mori *et al.*, 1999; Xue *et al.*, 1999). Beginning at the centre of a **seed voxel**, the algorithm moves in the direction of the principal diffusion orientation until reaching the boundary with another voxel, at which time the direction of the reconstructed tract changes to match the orientation of diffusion in the voxel it is entering. This process continues until a termination criterion is met, and is then repeated in the opposite direction from the seed point. It should be noted that the arrowheads shown at each voxel are present for the benefit of interpretation only—they have no physical significance, since diffusion orientation information is directionally nonspecific.

The differences between the early tractography algorithms are primarily in the choice of termination criteria and sampling policy. While FACT samples a trajectory direction exactly once per voxel, other approaches interpolate the original data to obtain local orientation information at a shorter scale, with the reconstruction typically involving short steps of a fixed distance (Basser *et al.*, 2000; Conturo *et al.*, 1999)—a strategy which results in smoother tract pathways than the FACT one. Meanwhile, anisotropy and tract curvature thresholds are commonly used as termination criteria for the reconstruction process, both of which help avoid tracking into

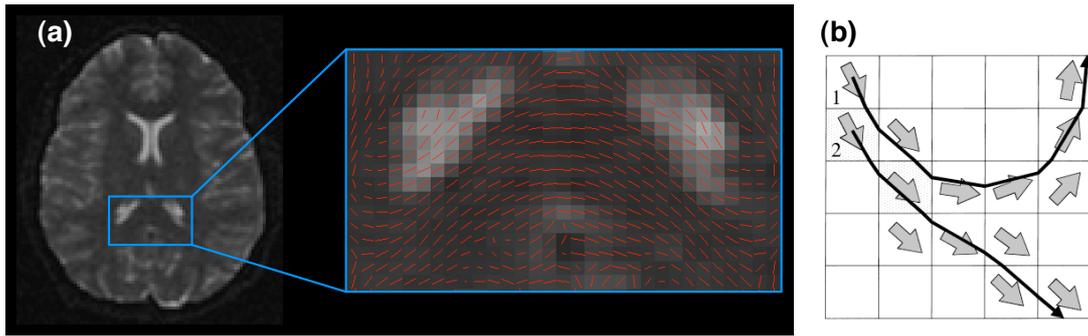


Figure 5.1: (a) Visualisation of the principal orientations and magnitudes of diffusion at each voxel in part of a dMRI image. (b) Reconstructing fibre pathways using the FACT algorithm. Subfigure (b) is reproduced from Xue *et al.* (1999).

grey matter or cerebrospinal fluid regions. Further discussion of these issues can be found in a review of fibre tracking methods by Mori & van Zijl (2002).

The simplest fibre tracking algorithms are completely deterministic—the principal eigenvector of the diffusion tensor is assumed to be a reliable and noise-free indicator of the local white matter trajectory. The problem, of course, is that the principal diffusion direction is neither of these things. Its reliability is never perfect, and will be affected by the number of gradient directions applied to the sample and any registration errors that occurred during the alignment of the component images, while noise is in fact omnipresent and will tend to “cause a computed trajectory to hop from tract to tract”, as Basser *et al.* (2000) have pointed out. Moreover, noise errors will accumulate as one moves further and further from the seed point. One way to try to circumvent this issue is to impose constraints on the tract reconstruction process which are informed by *a priori* knowledge about the geometry or topology of the underlying fasciculi (Conturo *et al.*, 1999; Poupon *et al.*, 2000). The benefit of these methods—as well as the extent of the problem that they attempt to tackle—is, however, difficult to predict in general terms, since the effects of noise (say) will depend on the particular protocol used to acquire the data, the shape of the tract, the signal-to-noise ratio, the anisotropy characteristics of the tissue, and so on. The final streamline itself gives no indication of the level of confidence that one can expect in the reconstruction.

More recently, streamline-based algorithms have been developed that attempt to indicate the variability that can result, when tracking from a single seed point, due to noise and uncertainty in the data. Some of these techniques are parametric, using a model to explain the data, while others are nonparametric, and therefore implicitly take any source of variability in the results into account. Fundamentally, however, all of these probabilistic approaches are based on the idea of replacing the single principal diffusion direction with a distribution over orientations, which indicates the uncertainty associated with the data at each voxel. One can then generate a family of streamlines from a single seed point using a Monte Carlo approach, sampling from these local distributions each time a new tracking direction is needed. Early work in this vein was published by Lazar & Alexander (2002) and Parker *et al.* (2003), who used the tensor shape to choose the variance of the orientation distributions. The approach demonstrated by Parker *et al.* was later developed further by Cook *et al.* (2004).

Bootstrap approaches to tractography are an example of a nonparametric statistical approach. Bootstrap is a resampling method, which requires that multiple measurements of the diffusion-weighted signal be taken for each diffusion gradient direction. Then, rather than using all of these data to fit a single diffusion tensor—which is the maximum likelihood approach taken by more simplistic algorithms—a subset of the data is sampled, with replacement, from the multiple measurements, and the tensor is calculated from this subset. A large number of these subsets are then extracted from this original data set, producing an empirical distribution over each of the free parameters in the diffusion tensor model. A general approach to using bootstrap to characterise uncertainty in dMRI data was put forward by Pajevic & Basser (2003),

Full data set				Sample 1	Sample 2
$A_1^{(1)}$	$A_1^{(2)}$...	$A_1^{(6)}$	$A_1^{(2)}$	$A_1^{(4)}$
$A_2^{(1)}$	$A_2^{(2)}$...	$A_2^{(6)}$	$A_2^{(5)}$	$A_2^{(4)}$
$A_3^{(1)}$	$A_3^{(2)}$...	$A_3^{(6)}$	$A_3^{(1)}$	$A_3^{(1)}$
$A_4^{(1)}$	$A_4^{(2)}$...	$A_4^{(6)}$	$A_4^{(2)}$	$A_4^{(6)}$
\vdots	\vdots		\vdots	\vdots	\vdots
regression ↓ \hat{D}				↓ $D^{(1)}$	↓ $D^{(2)}$

Table 5.1: Illustration of the application of bootstrapping to a dMRI data set containing repeated measurements. We denote the i th signal measurement using the k th gradient direction as $A_k^{(i)}$.

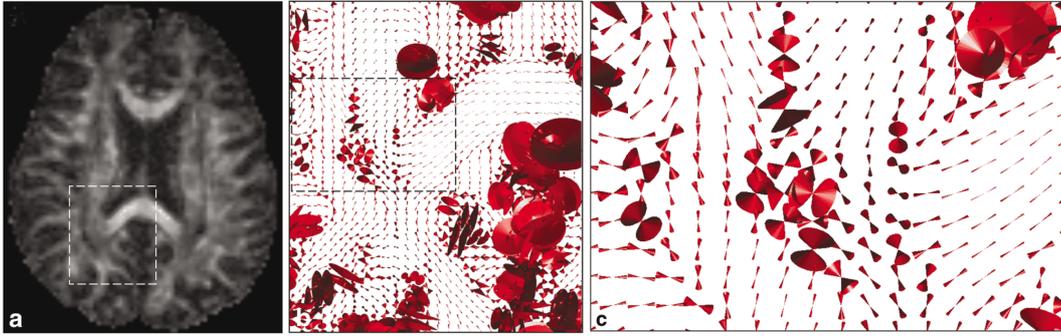


Figure 5.2: Orientation uncertainty in dMRI data, visualised as cones showing the 95% angular confidence interval at each voxel. Subfigure (b) corresponds to the area of (a) indicated with a box; likewise the further enlarged image (c). Reproduced from Jones (2003).

with applications to tractography following later (Jones & Pierpaoli, 2005; Lazar & Alexander, 2005).

Let us assume, for the sake of argument, that we have made six signal measurements for each of the gradient directions applied during a dMRI experiment. The diffusion tensor, D , can then be estimated from various subsets of these data, provided that at least the minimum six noncollinear gradient directions, plus a measurement with no diffusion weighting, contribute data to each subset. This is the principle employed by Jones (2003), and illustrated by Table 5.1. The maximum likelihood tensor is denoted by \hat{D} , while those estimated by sampled subsets of the data are denoted $D^{(1)}$ and so on. The latter can be used to estimate the uncertainty associated with the principal eigenvector, which is visualised in Fig. 5.2. Each set of sampled tensors for a given brain volume can then be used, in turn, to generate a single streamline from a chosen seed point, using a normal deterministic algorithm. The result will be a set of streamlines with a spatial distribution that reflects the variability encountered by the streamlining algorithm across the sample set—as shown in Fig. 5.3.

There are some interesting characteristics of the uncertainty elucidated in this way. Firstly, we can see by immediate inspection of Fig. 5.2 that the width of the 95% confidence interval on the principal diffusion orientation, which is depicted there, is highly variable between voxels. Near the middle of the corpus callosum splenium the confidence interval is extremely narrow. In this region, the maximum likelihood tensor would provide a reliable indication of the trajectory of this white matter structure. By contrast, the uncertainty is huge in areas which are composed primarily of *csf*—like near the bottom right of subfigure (b)—where

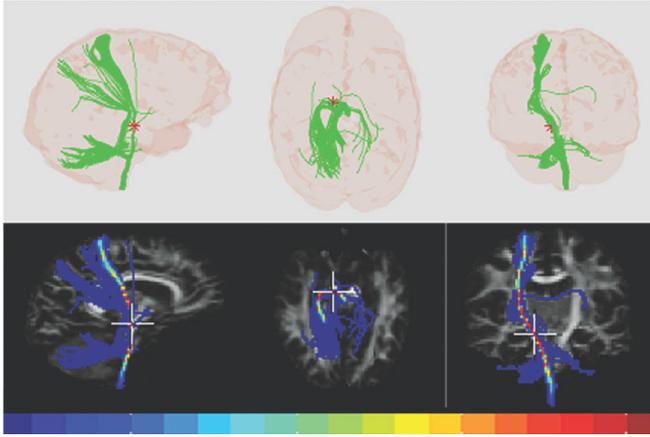


Figure 5.3: Results of applying the bootstrap method to tractography of the corticospinal tract. From a single seed point, which is indicated on each subfigure, a number of sample streamlines are produced (top). The proportion of the streamlines visiting each voxel can be counted to form a “visitation map” (bottom). Reproduced from Jones & Pierpaoli (2005).

diffusion is close to isotropic. Less predictable, however, is the effect of fibre crossings, which can be observed near the centre of (c). In this case, diffusion is approximately oblate, with two relatively large eigenvectors and one smaller one; and so the principal direction is less certain. The cone metaphor reflects this.

The necessity of acquiring multiple signal measurements for each diffusion gradient direction represents a problem for the basic bootstrap paradigm, because it will result in considerably extended scanning times without the improvement in angular resolution that would result from spending this time sampling more directions. Long scan times are particularly problematic in the clinical domain, since patients cannot be expected to remain still for long periods of time. Furthermore, the bootstrap method can substantially underestimate the degree of uncertainty in the tensor components when the number of repeated acquisitions is small (Chung *et al.*, 2006). However, a method known as the wild bootstrap offers to remove the need for multiple acquisitions when estimating the uncertainty in dMRI data (Whitcher *et al.*, 2007).

The wild bootstrap differs from “ordinary” bootstrapping in that it works with the residuals from a diffusion tensor fit to the signal data. If we describe a vector of unknown parameters, $\mathbf{x} = (D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}, \ln A_0)$ —where D_{xx} (and so on) are the tensor components, and A_0 is the signal without diffusion weighting—then the linear model used to estimate these components can be written out as

$$\mathbf{A} = \mathbf{B}\mathbf{x} + \boldsymbol{\epsilon},$$

where \mathbf{A} is a vector of observed log-signal values, \mathbf{B} is a matrix describing the diffusion gradient directions applied, and $\boldsymbol{\epsilon}$ is a vector of error terms. Thus we can evaluate an estimate for the parameters, $\hat{\mathbf{x}}$, using least-squares regression and our knowledge of \mathbf{A} and \mathbf{B} . As with other bootstrap methods, we do not need an explicit model for the errors, which are caused by noise and misregistration and so on. However, we subsequently use them to generate samples according to

$$A_k^{(i)} = \mathbf{B}_k \cdot \hat{\mathbf{x}} + h_k s_k^{(i)} \epsilon_k, \quad (5.1)$$

where \mathbf{B}_k and ϵ_k are the elements of \mathbf{B} and $\boldsymbol{\epsilon}$ corresponding to the k th direction, and $s_k^{(i)}$ has the simple probability mass function

$$\Pr(s_k^{(i)} = s) = \begin{cases} \frac{1}{2} & \text{for } s = \pm 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall k, i. \quad (5.2)$$

The constant h_k in Eq. (5.1) is used to ensure that the sampled residuals have the covariance structure required by the method (see Chung *et al.*, 2006, for details). Rather than repeatedly measure A_k , therefore, we instead resample the data by randomly permuting the signs of the residuals—i.e. by sampling from Eq. (5.2) for each value of i and k . Thus, only a single set of real measurements need be made, keeping scanning time short.

5.2 BEDPOST

It should be noted that the wild bootstrap introduces a dependence on the diffusion tensor model which is not present using ordinary bootstrap. Since the acquired data must be fitted to some kind of model for residuals to be available, the wild bootstrap is by nature a model-based resampling method. However, while the signal measured for each diffusion gradient applied is modelled using the diffusion tensor formalism, no model is used to explain the variability itself.

It is possible to go further, and model the observed data including their inherent uncertainty. This is the aim of another category of tractography algorithms, including the BEDPOST algorithm (Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques; see Behrens *et al.*, 2003b), which has been used for most of the practical parts of this thesis. By way of illustration of a fully model-based approach to tractography, and because of its centrality to work described later, this algorithm is fully described below.

The BEDPOST algorithm uses Markov chain Monte Carlo sampling to estimate diffusion MRI parameters. As above, the algorithm works with a vector of observed log-signal data, \mathbf{A} , and a model parameter vector, \mathbf{x} . However, because the diffusion tensor model can only usefully describe a single principal diffusion direction—since the second eigenvector is constrained to be orthogonal to the first—similar information can be embodied in a simpler model. In particular, Behrens *et al.* assume that the diffusion displacement distribution is a mixture of two Gaussians, in which one “compartment” is isotropic and the other is perfectly anisotropic, describing a single local tract orientation. The signal for the k th diffusion direction, μ_k , is then given by

$$\mu_k = A_0 \left((1-f) \exp(-b_k D) + f \exp(-b_k D \mathbf{G}_k^T \mathbf{R} \mathbf{M} \mathbf{R}^T \mathbf{G}_k) \right), \quad (5.3)$$

where b_k is the k th scalar b -value, \mathbf{G}_k is the k th diffusion encoding direction represented as a column vector,

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and \mathbf{R} rotates \mathbf{M} to align with the fibre direction in the voxel, which requires two implicit angles (θ, ϕ). Compare Eq. (5.3) with the standard dMRI formulations in Eqs (4.7) and (4.8). The natural index of anisotropy arising from this model is the mixture coefficient, f , which we refer to as the **anisotropic volume fraction** (AVF). Note that the model provides no information about anisotropy perpendicular to the direction encoded by \mathbf{R} ; but then such information is not directly relevant to streamline tractography.

Under the generative model of local diffusion described by Eq. (5.3) and the assumption that noise is independent and identically distributed for each measurement, the likelihood of the observed data is given by

$$P(\mathbf{A}|\mathbf{x}) = \prod_k P(A_k|\mathbf{x}), \quad (5.4)$$

where

$$P(A_k|\mathbf{x}) \sim N(\mu_k, \sigma^2); \quad (5.5)$$

and so the full parameter vector is $\mathbf{x} = (A_0, D, f, \theta, \phi, \sigma)$. The posterior distribution over these parameters is given by Bayes’ rule:

$$P(\mathbf{x}|\mathbf{A}) = \frac{P(\mathbf{A}|\mathbf{x})P(\mathbf{x})}{\int P(\mathbf{A}|\mathbf{x}')P(\mathbf{x}')d\mathbf{x}'}. \quad (5.6)$$

For the purposes of fibre tracking, however, the most important parameters at each voxel are the angles which provide tract orientation information. If we wish to obtain distributions over $\mathbf{x}_1 = (\theta, \phi)$, we will need to calculate the marginal distribution given by

$$P(\mathbf{x}_1|\mathbf{A}) = \int P(\mathbf{x}|\mathbf{A})d\mathbf{x}_2, \quad (5.7)$$

where $\mathbf{x}_2 = (A_0, D, f, \sigma)$, a vector consisting of the remaining parameters. Both the evidence term in the denominator of Eq. (5.6) and the marginal distribution of Eq. (5.7) require the evaluation of complex integrals, however, and cannot be expected to be soluble analytically. We therefore turn to MCMC sampling to evaluate them empirically.

The priors, $P(\mathbf{x})$, in Eq. (5.6) are chosen by the authors to be uninformative, except where ensuring positivity is appropriate: in A_0 and f . Initialisation for the Markov chains is provided by performing a normal least-squares diffusion tensor fit to the data at each voxel, and using tensor analogues of each parameter. Samples for σ are generated using a Gibbs sampler, and all other parameters are sampled using the Metropolis–Hastings algorithm. Proposal distributions for the latter are zero-mean Gaussians whose variance is tuned to maintain an acceptance rate of 0.5.

The generative model for the noisy data, Eq. (5.5), takes the form of a normal distribution with known mean—given knowledge of the partial parameter vector $\mathbf{x}_3 = (A_0, D, f, \theta, \phi)$ —and unknown variance. This is a common and therefore well-characterised situation. Using a gamma prior distribution for the precision, $\tau = 1/\sigma^2$, viz.

$$P(\tau|\alpha, \beta) = \text{Gamma}(\alpha, \beta) = \frac{\tau^{\alpha-1} \beta^\alpha e^{-\beta\tau}}{\Gamma(\alpha)},$$

where $\Gamma(\cdot)$ is the gamma function, the posterior over τ given data \mathbf{A} is another gamma distribution:

$$P(\tau|\alpha, \beta, \mathbf{A}, \mathbf{x}_3) = \text{Gamma}\left(\alpha + \frac{K}{2}, \beta + \frac{1}{2} \sum_k (A_k - \mu_k)\right),$$

where K is the total number of gradient directions acquired. This is used by *BEDPOST* as the conditional distribution for the Gibbs sampler, although the authors do not explicitly state how they chose the prior hyperparameters α and β .

The marginal distribution for \mathbf{x}_1 is trivially extracted from the samples over \mathbf{x} by considering only θ and ϕ from each sample vector. The tractography part of the algorithm—which the authors call *ProbTrack*—then uses these samples to reconstruct a set of “probabilistic streamlines” using a normal streamlining approach. Given a seed voxel, \mathbf{a} , the process is as follows.

1. Start with the current “front” of the streamline set to \mathbf{a} .
2. Select a random sample, (θ, ϕ) , from $P(\theta, \phi|\mathbf{A})$ at the streamline front.
3. Move the front some small distance in the direction of (θ, ϕ) .
4. Return to step 2, and repeat until a stopping criterion is met.

The stopping criteria are not strict, stipulating only that a streamline is not allowed to curve by more than about 80° , and that a streamline will be terminated if it leaves the brain or enters an area that it has already visited.

To evaluate the direction of propagation at any location in the brain, not just those that coincide with voxel centre points, some kind of interpolation scheme is required. The authors use a probabilistic analogue of trilinear interpolation, in which a sample is drawn from one of the two adjacent voxels in each dimension according to how close the sample location is to each of them. Indexing in voxel steps, the sample location is taken from the p.m.f.

$$\Pr(x = v) = \begin{cases} \text{ceil}(x) - x & \text{for } v = \text{floor}(x) \\ x - \text{floor}(x) & \text{for } v = \text{ceil}(x), \end{cases}$$

where *floor* and *ceil* are the usual floor and ceiling functions. If $x = \text{floor}(x)$ —that is, x falls exactly on a voxel location—then the sample is taken from that voxel with unit probability.

This procedure for generating streamlines is repeated a large number of times (typically 5000) for a particular seed point, generating a spatial distribution for the tract running through the seed point at \mathbf{a} . This distribution may be usefully discretised by counting up the number of streamlines passing through each voxel and associating this count with the voxel volume. An

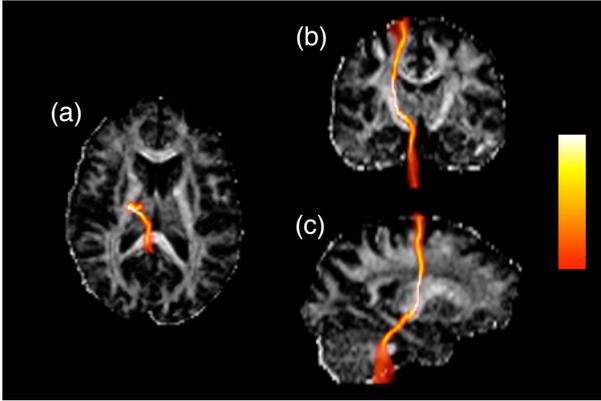


Figure 5.4: Example discretised spatial distribution from the BEDPOST/ProbTrack tractography algorithm, showing the corticospinal tract in axial (a), coronal (b) and sagittal (c) maximum intensity projections. The underlying greyscale image shows AVF in the slice in-plane with the seed point in each case. White indicates that nearly all streamlines pass through the local voxel, while red means that very few do. The full colour scale is shown.

example of the result is shown in Fig. 5.4. These data can be interpreted as confidence bounds on the location of the most probable tract passing through the seed point.

Behrens *et al.* showed, in their paper, that the levels of uncertainty estimated by their method are comparable with those estimated by the bootstrap approach described by Jones (2003)—thus justifying, to some extent, the additional assumptions that they make in their fully model-based approach. The advantage of this added model specificity, meanwhile, is an improved sensitivity.

A standard implementation of BEDPOST is freely available as part of the FSL package of software tools (Smith *et al.*, 2004), which is written and maintained by the FMRI centre at the University of Oxford.

A number of variations on, and extensions of, the BEDPOST method have been proposed. Friman *et al.* (2006) describe another alternative model for diffusion at a voxel, which is essentially the tensor model, but with the two smaller eigenvalues constrained to be equal—that is, $\lambda_2 = \lambda_3 = \alpha$ —thereby producing the form

$$\mu_k = A_0 \exp(-\alpha b_k) \exp\left(-(\lambda_1 - \alpha) b_k \mathbf{G}_k^T \mathbf{RMR}^T \mathbf{G}_k\right). \quad (5.8)$$

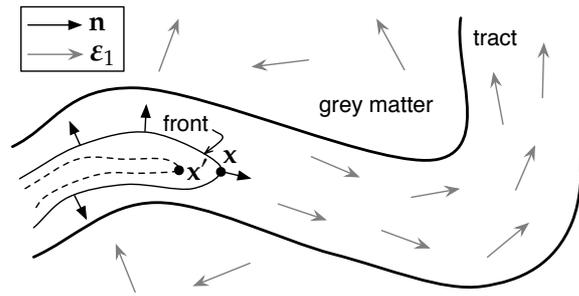
The authors also use a more theoretically justified noise model, whose variance depends on the signal value; and they use point estimates for the “nuisance” parameters in the model in order to reduce its computational demands.

Neither the compartment model described by Eq. (5.3) nor the constrained model of Eq. (5.8) can account for more than one fibre orientation at a voxel. Rather, multiple fibre orientation information is manifested as increased uncertainty in the single orientation that they can represent. However, both models can be generalised to handle this case, which occurs commonly in the brain at typical imaging resolution—the compartment model by adding extra anisotropic compartments (Behrens *et al.*, 2007), and the constrained case by modelling additional tensors (Hosey *et al.*, 2005). It is generally wise to use the simplest model that explains the data satisfactorily at each voxel, rather than simply to fit multiple fibre orientations at every location in the brain. Hosey *et al.* achieve this by fitting one and two tensor models at each voxel and using probabilistic model selection to choose between the results, while Behrens *et al.* fit a single, complex model but apply a technique known as automatic relevance determination (see MacKay, 1995, §7) to factor out unneeded parameters.

5.3 Fast marching

Streamline generation is not the basis for all fibre tracking algorithms; although it is, as we have mentioned, the most common. One alternative general approach is to propagate a 3-D surface or front in all directions from the seed point at once, such that its speed is faster in some directions than in others—a method called *fast marching tractography* (FMT; see Parker *et al.*, 2002b). A speed function is used to define how fast the front moves as it progresses through

Figure 5.5: Front propagation in fast marching tractography. The speed function, F , is designed so that the front of the spreading region grows most quickly where its normal vector, \mathbf{n} , aligns closely with the principal eigenvector of the local diffusion tensor, ϵ_1 . Thus the front will move fastest along paths with smoothly varying principal diffusion orientation. Arrowheads on the eigenvectors are notional. After Parker *et al.* (2002b).



the brain. Parker *et al.* (2002a) use the speed function

$$F(\mathbf{x}) = \min \{ |\epsilon_1(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})|, |\epsilon_1(\mathbf{x}') \cdot \mathbf{n}(\mathbf{x})|, |\epsilon_1(\mathbf{x}) \cdot \epsilon_1(\mathbf{x}')| \}, \quad (5.9)$$

where $\mathbf{n}(\mathbf{x})$ is the local normal to the front at point \mathbf{x} , and $\epsilon_1(\mathbf{x})$ is the first eigenvector of the local diffusion tensor. The point \mathbf{x}' represents the position of a neighbouring voxel that has already been passed by the front. These terms are visualised in Fig. 5.5.

As we follow this propagating front out from the seed point, we can establish a “time of arrival” for each voxel in the brain. Wherever the front moves fastest, the time of arrival to voxels along its route will be low. One can do target-based tractography by then performing a gradient descent in a time of arrival map, from the target voxel back to the seed. Exploratory tractography from a seed point is also possible by using every other voxel in the brain as a target point in turn, and retaining those pathways which are most plausible under some criterion, such as the minimum or average value of the speed function along them.

The performance of FMT hinges on the choice of speed function. Parker *et al.* (2002b) discuss alternative forms for the speed function, although they limit themselves to the case where the first tensor eigenvector can be considered a reliable indicator of tract direction. Since then, however, Staempfli *et al.* (2006) have described a set of four speed functions, from among which their FMT algorithm selects, depending on the tensor shape at \mathbf{x} and \mathbf{x}' . This allows their method to track through regions in which diffusion has an oblate, rather than prolate, profile.

5.4 High angular resolution methods

A number of models of diffusion have been developed for the purpose of elucidating the orientations of multiple fibre populations within a voxel. Some of these are direct extensions of simpler models, as we have already seen, while others were designed from the outset to work with crossing fibres.

The need for more complex models than the tensor model in tractography has been touched upon earlier in this chapter, but Fig. 5.6 demonstrates the issue explicitly (see also Frank, 2001). With a single fibre orientation per voxel, the tensor model is an adequate model, effectively representing the diffusion profile expected for this case, as in subfigure (a). On the other hand, we would like to be able to recover a profile encapsulating two fibre orientations when this is justified (b), but instead the tensor can only represent a directionally nonspecific profile (c). In order to track effectively through regions of crossing fibres, however, the structure in the inherent diffusion profile must be retained.

The first requirement for successful elucidation of crossing fibre architectures is, then, a diffusion model that is capable of representing their relatively complex structure; but there are also commensurate acquisition requirements. Since more complex models have more parameters, and in particular because they aim to more fully represent the diffusion profile, larger numbers of gradients must be applied to improve the angular resolution of the scan. For this reason, the modelling and acquisition techniques that aim to represent complex intravoxel architectures are called high angular resolution diffusion imaging (HARDI) methods. Secondly, in order to produce strong enough contrast between the signal effects of each fibre population, greater diffusion weighting—corresponding to greater values of b , the weighting coefficient—is usually applied. This can be achieved by increasing gradient strength or diffusion time.

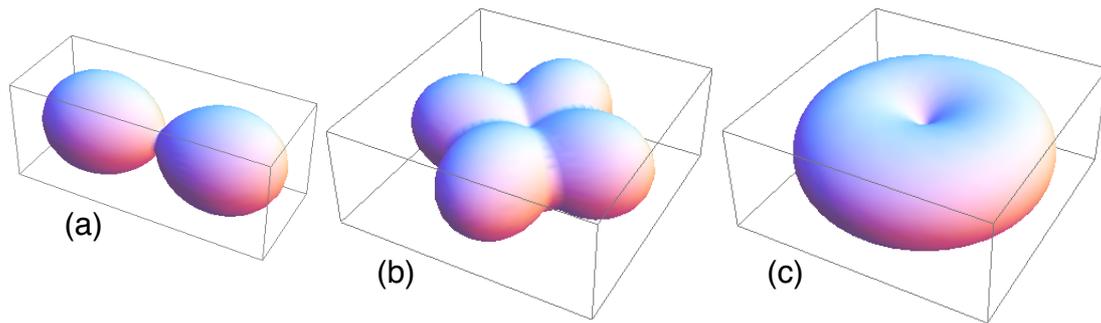


Figure 5.6: Visualisation of diffusivity as a function of gradient orientation. When there is a single fibre population within a voxel, it produces a diffusivity profile like the one in (a), which is well represented by the diffusion tensor model. In the presence of two orthogonal populations, the true profile is something like (b), but a single diffusion tensor is only capable of representing the ambiguous case shown in (c).

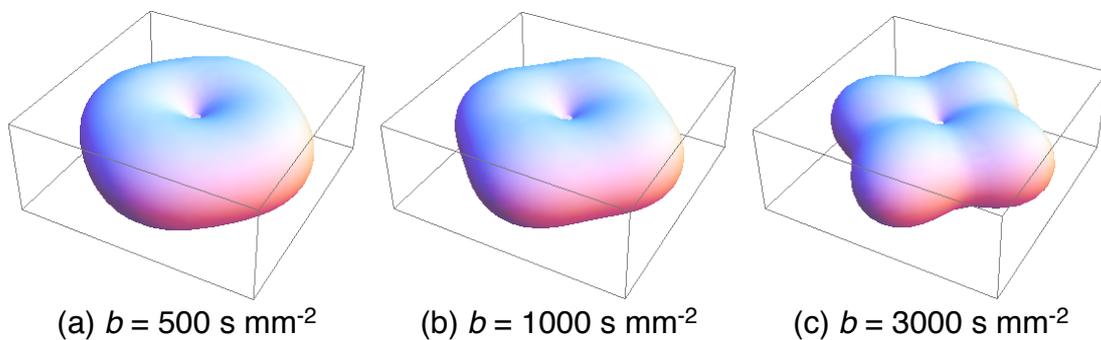


Figure 5.7: Dependence of the signal on b -value. Multiple fibre orientations are better contrasted at higher levels of diffusion weighting. Peak diffusivity in each of the two component tensors was $7.5 \times 10^{-4} \text{ mm}^2 \text{ s}^{-1}$.

The effect on angular contrast of increasing the b -value is shown in Fig. 5.7, and described in Alexander *et al.* (2001). It should be borne in mind that unfortunately, higher b -values also produce less overall signal—since the weighting factor determines the level of attenuation in the signal due to diffusion effects—so the signal-to-noise ratio of the acquisition is lower.

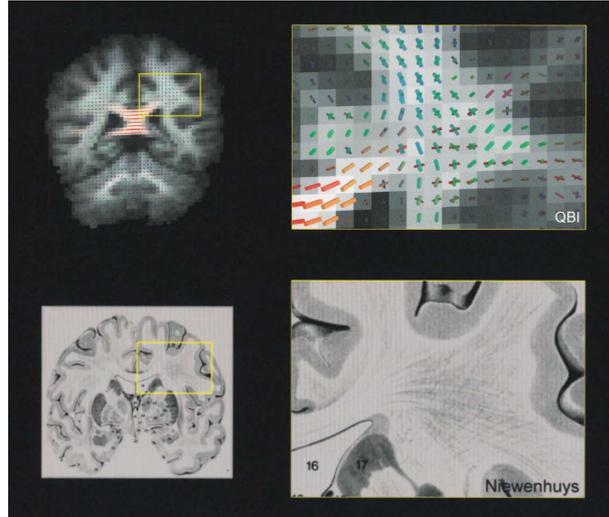
One way to handle multiple fibre directions is to use multiple tensors (Tuch *et al.*, 2002)—an approach we have already seen employed, in a constrained form, by Hosey *et al.* (2005). Under this model, the diffusion displacement distribution is assumed to be a mixture of Gaussians with different covariance structures. Two tensors are able to faithfully represent the situation shown in Fig. 5.6(b), although a third tensor would need to be used for the case of three fibre populations, and it is often not possible to know *a priori* how many fibre populations are expected within a given voxel.

It should be noted that a multiple tensor model assumes that the diffusing water molecules do not move between fibre populations during the course of the experiment. This is known as the assumption of *slow exchange*, and it is a typical assumption in the analysis of crossing fibre structure. Since the root-mean-squared diffusion distance for a typical dMRI protocol is of the order of $10 \mu\text{m}$, compared to a typical axon diameter of a few microns, this assumption is thought to be a reasonable one for most purposes. Fibre tracts consist of bundles of hundreds of axons, and so diffusion over the width of a few axons will rarely exchange between bundles.

5.4.1 Using q -space

An alternative general approach to the crossing fibre problem is to employ q -space diffusion imaging. As we saw in §4.3, q -space imaging allows us to recover an arbitrary displacement

Figure 5.8: Results from \mathbf{q} -ball imaging and comparison with invasive tracing. The \mathbf{q} -ball reconstruction of fibre orientations effectively represents the fanning out of pathways emerging from part of the corpus callosum, which can be seen in an atlas of the central nervous system (Niewenhuys, 1996). Reproduced from Tuch *et al.* (2003).



distribution in a model-free manner, by taking a Fourier transform of MRI signal information acquired using an appropriate scheme. A scheme suitable for recovering crossing fibre orientations was described by Wedeen *et al.* (2005), using 515 \mathbf{q} -vectors and a maximal b -value equivalent of $17,000 \text{ s mm}^{-2}$. Having recovered a spatial displacement distribution, $P(\mathbf{r})$, an *orientation distribution function* (ODF) can be calculated by projecting the distribution onto the unit sphere. That is,

$$\Psi(\hat{\mathbf{r}}) = \int_0^\infty P(\rho\hat{\mathbf{r}})\rho^2 d\rho, \quad (5.10)$$

where $\mathbf{r} = \rho\hat{\mathbf{r}}$. In this case the authors use the squared vector length, ρ^2 , as a weighting factor. The ODF then provides the information needed to perform tractography, using a streamline method or otherwise, in the region. (It should be noted, however, that the ODF has no probabilistic interpretation because it is not properly normalised.) This approach is called *diffusion spectrum imaging* (DSI).

The biggest problem with DSI is its acquisition requirements. The protocol makes no real attempt to satisfy the narrow gradient pulse assumption, so the demands it makes on gradient hardware are not extreme; but because it samples \mathbf{q} -space quite thoroughly, imaging a brain volume at a reasonable resolution takes far longer than a comparable DTI protocol.

A step towards reduction of the \mathbf{q} -space sampling requirements of DSI was taken by the development of so-called \mathbf{q} -ball imaging (Tuch *et al.*, 2003; Tuch, 2004). In this case, the length of the sampled \mathbf{q} -vectors is fixed so that they lie on a sphere. The authors show that an ODF can then be recovered directly by means of an integral transform called the Funk–Radon transform, which has its roots in computed tomography, a medical imaging technique using x-rays. The authors also describe a method for calculating this transform that is reasonably simple and computationally inexpensive. Fast marching tractography has since been demonstrated using the \mathbf{q} -ball ODF as a speed function (Campbell *et al.*, 2005).

It has been shown that the \mathbf{q} -ball method produces ODF information that is in fairly good agreement with standard, invasive tracing work (see Fig. 5.8), and certainly provides more useful information for tractography in crossing fibre regions than the tensor model (Fig. 5.9). Note that in Fig. 5.9(a) the first tensor eigenvector represents a more or less arbitrary orientation in regions where the crossing occurs. This is consistent with the degenerate representation expected under this model (cf. Fig. 5.6).

Jansons & Alexander (2003) describe an alternative to an orientation distribution function called *persistent angular structure* (PAS). As with the ODF formulation, the aim is to capture the orientation information in the signal which is important for tractography, whilst discarding the less salient radial information. The radial part of the diffusion displacement distribution is

therefore factored out and represented by a Dirac δ -function, viz.

$$P(\mathbf{r}) = \frac{p(\hat{\mathbf{r}}) \delta(|\mathbf{r}| - \rho)}{\rho^2},$$

where the function $p(\hat{\mathbf{r}})$ is the P_{AS} , the angular component of the distribution. Here, ρ is a parameter that has to be chosen independently. By means of an optimisation which is constrained by the relationship between the data and the displacement distribution—embodied in Eq. (4.13)—the authors arrive at the maximum entropy solution

$$p(\hat{\mathbf{r}}) = \exp \left(\lambda_0 + \sum_j \lambda_j \exp(i\mathbf{q}_j \cdot \rho \hat{\mathbf{r}}) \right), \quad (5.11)$$

where $\{\lambda_j\}$ are constants to be found. The maximum entropy solution is the most uninformative function possible, subject to the constraints imposed by the data. The intuition of this approach is to encode just the angular structure “reported” by the acquired data, without introducing extra information by making additional assumptions.

A big—perhaps the biggest—advantage of P_{AS} -MRI is its modest acquisition requirements. Jansons & Alexander use a scheme involving just 54 nonzero \mathbf{q} -vectors, compared to hundreds for a typical DSI or \mathbf{q} -ball experiment. The trade-off, however, comes in computation time. Since the P_{AS} , Eq. (5.11), is a nonlinear combination of functions, reconstruction times for P_{AS} -MRI are typically orders of magnitude longer than those required by the other, linear techniques. With present computing power, the time needed to fully process a large data set could be prohibitive.

5.4.2 Spherical deconvolution

A further subcategory of $HARDI$ methods use a technique called spherical deconvolution, which allows one to recover an ODF without relying on the Fourier relationship between the $dMRI$ signal and the displacement distribution, which is anyway only approximate since the narrow gradient assumption is not fulfilled. Instead, the fundamental assumption here is that the signal arises from the convolution of an ODF with a “response function”, which is assumed to be invariant across all white matter in the brain, with partial volume effects accounting for all nonorientational variability (Tournier *et al.*, 2004). Slow exchange is also assumed. We therefore write, for a particular b -value,

$$A(\theta, \phi) = \sum_i f_i \mathbf{R}_i S(\theta) = \Psi(\theta, \phi) \otimes S(\theta), \quad (5.12)$$

where θ represents the polar angle and ϕ the azimuthal angle in spherical polar coordinates, f_i is the volume fraction of the i th fibre population, and \mathbf{R}_i is a rotation matrix representing its orientation. The symbol \otimes represents convolution on the unit sphere. We note that the unit vector $\hat{\mathbf{r}}$ used above as the ODF parameter is related to the two angles by

$$\hat{\mathbf{r}} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta).$$

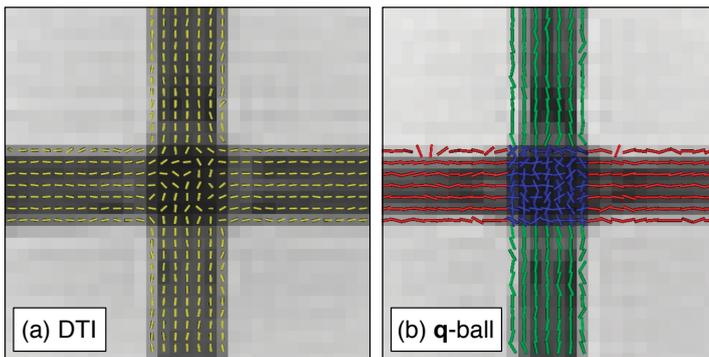
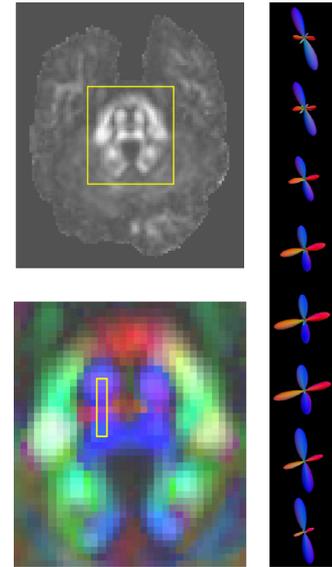


Figure 5.9: Fibre orientation information reconstructed using DTI and \mathbf{q} -ball methods for a specially constructed phantom, mimicking orthogonal crossing fibre populations. Adapted from Perrin *et al.* (2005).

Figure 5.10: Results of applying the spherical deconvolution method in a region of fibre crossing in the pons. The two fibre orientations appear quite distinct from one another and are qualitatively accurate representations of the underlying architecture. Reproduced from Tournier *et al.* (2004).



The response function, S , is a function only of θ because it is taken to be axially symmetric. Given the knowledge of an appropriate response function, the ODF can therefore be deconvolved out of the signal profile, at least in principle.

It is worth noting that the spherical deconvolution model is related to the anisotropic component of the BEDPOST partial volume model, Eq. (5.3). In that case the response function represents Gaussian diffusion along a single orientation, and the ODF is a δ -function whose orientation corresponds to that of the modelled fibre pathway. PAS-MRI can also be framed as a deconvolution (Alexander, 2005).

Tournier *et al.*, by contrast, use an ODF which can represent multiple directions; and they represent it, along with the response function, in terms of a set of functions known as the spherical harmonics (Riley *et al.*, 2002). These functions form an orthonormal basis set over the sphere, and their use in general spherical deconvolution problems has been described by Healy *et al.* (1998). Representation of the signal profile, A , using these basis functions had already been described (Alexander *et al.*, 2002; Frank, 2002). Under this parameterisation, the ODF can be recovered by means of a straightforward set of matrix multiplications, given knowledge of the response function—which the authors establish by observing the signal profile in strongly anisotropic parts of the brain.

The method has been demonstrated to work well for resolving fibre crossings in simulations and in real data acquired with a modest 60 gradient directions at a b -value of about 3000 s mm^{-2} (see Fig. 5.10). The authors estimate that using these acquisition parameters, two fibre orientations with a separation of 60° can be recovered with a standard deviation of around 9° . The minimum resolvable separation is estimated to be about 40° .

The validity of the assumption of equivalent response throughout the brain is hard to establish, but the most significant shortcoming of the method is probably its sensitivity to artefacts caused by noise. Recently developed methods for regularising the ODF (Sakaie & Lowe, 2007; Tournier *et al.*, 2007) promise to mitigate this issue significantly, however—even when the signal-to-noise ratio is low. Thus it may be possible to apply the method to recover useful orientation information even at the lower b -values commonly used in DTI experiments. A parametric version of the spherical deconvolution method has also been recently developed, allowing Bayesian statistics to be used to infer an ODF (Kaden *et al.*, 2007).

5.5 Applications and challenges

Given the increasingly formidable array of ideas and innovations which have been thrown at the fibre tracking problem, it is natural to ask what scientific uses there may be for reliable

tractography methods once they have been developed. At present, there are two general categories of application for these algorithms which have appeared in the literature.

The first application might be loosely described as connectivity analysis. Despite the fact that tractography is still very much a field in its infancy, it is already beginning to provide information about the brain's internal connections which are corroborating the findings of more well established—and more invasive—neuroscientific techniques. In an impressive piece of work, Behrens *et al.* (2003a) demonstrated, using tractography, that voxels in the thalamus can be effectively categorised by the targets of their most likely projections into cortex. The resulting thalamic parcellations are in close agreement with atlas data (see Fig. 5.11), and have been further reinforced by functional MRI results (Johansen-Berg *et al.*, 2005). Similar principles have been applied to the corpus callosum (Huang *et al.*, 2005), and used to identify boundaries between cortical regions based on their connectivity (Johansen-Berg *et al.*, 2004).

The second category of application encompasses the segmentation and visualisation of specific tracts. The emphasis in this case is more clinical than neuroscientific, since segmenting a particular tract is often a precursor to comparative analysis of anisotropy—or some other indicator of pathology—between a patient group and controls. We will not expand further on the segmentation application here, however, because it will be the focus of the next three chapters; and therefore will be described fully elsewhere. Tract visualisation can be useful in its own right as a preoperative surgical planning tool, since any invasive treatment will naturally try to minimise damage to important connective pathways—although at present it is highly advisable to avoid setting too much store by tractographic results in such critical applications (Kinoshita *et al.*, 2005).

Notwithstanding their increasing popularity and promising early results, tractography methods have some outstanding theoretical and practical limitations. The problem of handling crossing fibres cannot be said to be fully solved, especially in the relatively high noise and low angular resolution regime which is common in clinical MRI scanning. There is also an additional degeneracy which is widely recognised, but whose impact has not yet been fully characterised: the problem of “kissing” fibres (Basser *et al.*, 2000). From a fibre tracking point of view, it is important to be able to distinguish the two intravoxel architectures shown in Fig. 5.12, but a recovered ODF will usually not provide enough information to do so.

However plausible the reconstructed tracts may appear to be, the issue of validation is a significant one. Efforts to validate tractography methods have recently increased, and include computer simulation work (as in Hosey *et al.*, 2005) and studies with physical phantoms designed to mimic biological white matter (Campbell *et al.*, 2005; Perrin *et al.*, 2005). In addition, Bürgel *et al.* (2006) have generated maps of the routes of a number of fasciculi, based on postmortem histology, for comparison with tractography results. We saw, in Fig. 5.8, evidence of qualitative agreement between the *q*-ball ODF and fasciculus crossing information derived invasively; and in a similar way Schmammann *et al.* (2007) demonstrated a very respectable agreement between DSI-based tractography and histology in the monkey brain. The matches are still far from perfect, however; and are often demonstrated in rather idealised conditions. The DSI scan used by Schmammann *et al.*, for example, was performed on a 4.7 T system and

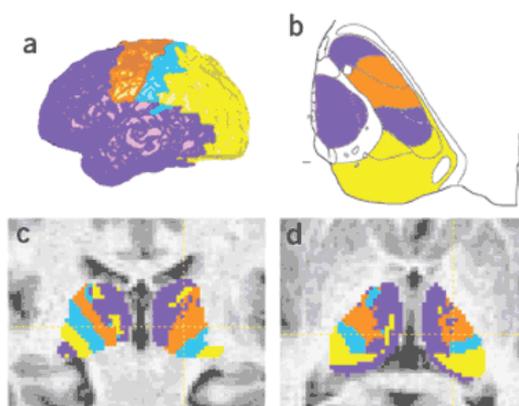
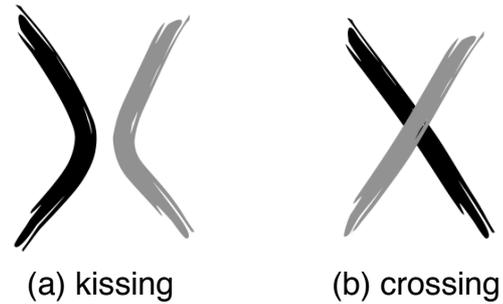


Figure 5.11: Thalamic parcellation using probabilistic tractography. Dividing the human brain into major cortical regions (a) and colour coding thalamic voxels according to their most probable projection into cortex using tractography (c,d) yields results broadly in good agreement with information obtained using invasive methods (b). Reproduced from Behrens *et al.* (2003a).

Figure 5.12: Kissing and crossing fibre architectures. In **(a)** the two fibre populations bend away from one another, whilst in **(b)** they cross or interdigitate. Since the angular information intrinsic to each of these scenarios is very similar, it is hard to tell them apart from their ODFs.



took 25 hours to complete. Such protocols are clearly useless to the clinician.

5.6 Summary

In this chapter we have attempted to provide a sense of the spectrum of extant approaches to the fibre tracking problem. We have focussed on giving a sense of the breadth of the alternatives, to avoid provoking informational indigestion in the reader (or the author), and have therefore omitted one or two notable techniques due to their similarity to other methods. It should be evident that the range of proposed solutions is wide, although they differ with respect to a fairly small number of core principles. Streamline-based tracking methods are the most widespread, but the model of orientation density is an important factor. There is, as yet, no clear reason to use one particular technique over all others, and studies based on tractography would be well advised to justify their choice of algorithm according to the nature of their aims.