

Parametric Models for Complex-Valued HARDI Data

Sofia C. Olhede^{a,*}, Gareth J. Barker^b, Rexford D. Newbould^c, Brandon
Whitcher^c

^a*Department of Statistical Science, University College, London, United Kingdom*

^b*King's College London, Institute of Psychiatry, Department of Clinical Neuroscience,
Centre for Neuroimaging Sciences, London, United Kingdom*

^c*Clinical Imaging Centre, GlaxoSmithKline, London, United Kingdom*

Abstract

Diffusion-weighted magnetic resonance imaging (DW-MRI) data are the noisy observations of the Fourier transform (FT) from a diffusion probability density function (PDF). If the diffusion PDF is symmetric in space then the DW-MRI measurements are (theoretically) real-valued; while if the PDF exhibits some asymmetry in space then the DW-MRI measurements are (theoretically) Hermitian complex-valued. A model for the complex-valued FT of the diffusion PDF is introduced in this paper from the three-dimensional Boltzmann equation, suitable for complex-valued high angular resolution diffusion imaging measurements. This model, corresponding to the characteristic function of the PDF, may be expanded in terms of cumulants of the PDF that are functions of its moments. The equation therefore characterizes the PDF in terms of its moments, where the latter can be directly related to physical characteristics of the diffusion PDF. The structure of the PDF of an idealized bending or forking fibre we argue that bending fibers produce “near-linear” phase functions over a range of q -values, the sphere of diffusion encoding directions, without corresponding to net transport. **Gareth = This sentence got lost somewhere in the middle!** To estimate such complex-valued structure from HARDI observations, spurious phase is first removed from the data via a pre-processing step. Voxels that are not Hermitian symmetric are removed from subsequent analysis. For such voxels, for which the pre-processing was successful, a locally linear phase is estimated using least squares, an estimate from which we argue the axes of locally lin-

*Corresponding author.

Email address: s.olhede@ucl.ac.uk (Sofia C. Olhede)

ear asymmetry in the diffusion PDF may be determined if the Hermitian structure differs significantly from the noise. We propose analytic asymmetric models for PDFs, discuss bending and forking in detail, and show the performance of our method on simulated and *in vivo* data acquired from a typical clinical setting.

Keywords:

DTI, phase, diffusion, brain, non-Gaussian.

1. Introduction

High angular resolution diffusion imaging (HARDI) data has had a great impact on the understanding of complex white-matter microstructure. From HARDI data it is possible to estimate properties of the local spatial diffusion of water molecules via the diffusion probability density function (PDF). HARDI data corresponds to multiple images acquired with different diffusion encoding directions at a high b -value, as well as at a low or zero b value, and this collection of images may be used to deduce orientational structure in the spatial domain. The b -value quantifies the influence of the diffusion-encoding gradients on the acquired images, and is a function of the strength of the gradient pulses, their duration as well as the time between pulses. HARDI data are often collected at a higher b -value than is normal for diffusion tensor imaging (DTI), to enhance any directional structure although collecting data at higher values of b reduces the SNR (signal-to-noise ratio) of the observations. HARDI observations are Fourier domain information and must be converted into spatial information, usually by an approximation to an inverse Fourier transform (IFT). To be able to make the necessary inversion, additional constraints must be placed on the observed white-matter structure, as the number of observations at each voxel fall short of the number needed for a numerically stable IFT procedure.

A number of methods exist for the analysis of orientational structure, with the most straightforward model corresponding to diffusion tensor imaging (DTI) (Basser et al., 1994a,b), where a Gaussian PDF is fitted at each voxel. There is an analytic expression for the IFT of the characteristic function (Fourier transform of a PDF) of a Gaussian PDF, so the estimated parameters in q -space have a direct spatial interpretation. There are limitations to the DTI model, and to be applicable across more heterogeneous structures either a mixture of Gaussian PDFs must be fitted (Alexander, 2005)

or non-parametric methods must be used. Popular non-parametric methods include the approach of Jansons and Alexander (2003), corresponding to a non-parametric estimation method of the orientational structure associated with a single fixed spatial scale. Tuch (2004) proposed fully non-parametric method of estimating spatial orientational structure using the Funk-Radon Transform (FRT). Improvements to the implementation of the FRT have been proposed by Hess et al. (2006) and Descoteaux et al. (2007). These are all methods of estimating an orientational distribution function (ODF). By an ODF we refer to any voxel-specific representation of spatial orientational structure, whether corresponding to a weighted average of structure across spatial radii, or associated with a fixed radius.

It is also possible to describe more complicated heterogeneous structure of the diffusion PDF non-parametrically using generalized diffusion tensors by expanding the logarithm of the characteristic function in terms of an infinite number of higher-order terms (Özarslan and Mareci, 2003; Liu et al., 2005a, 2004). Care must be taken to appropriately normalize the higher-order moments by b Liu et al. (2005a). Practical application of such ideas corresponds to diffusion kurtosis imaging (DKI), where the non-Gaussian higher-order structure of the diffusion PDF is quantified (Jensen et al., 2005; Lu et al., 2006; Hui et al., 2008). In DKI a tensor is calculated using higher-order terms in the Kramer-Moyal expansion of the three-dimensional Boltzmann equation, and yields additional information about the directional structure of the diffusion PDF. In Lazar et al. (2008) an ODF was calculated directly from the estimated diffusion tensor and diffusion kurtosis values. The higher-order terms are necessary to resolve low crossing angles and provide information about the high-frequency components of the ODF (Lazar, 2010).

All of the methods previously discussed focus on magnitude MRI data, and ignore the phase information from the inherently complex-valued MR acquisitions. It has been proposed that movement, susceptibility effects, and other random offsets to the phase, diminish any advantage in analysing the complex-valued observations directly. Notable exceptions include Liu et al. (2005a), Aksoy et al. (2008) and Newbould et al. (2008). Arguments have been made for the case that DW-MRI measurements are inherently real-valued and positive (Wedeen et al., 2005), thus eliminating the need to study the complex-valued MRI data other than, for example, the goal of noise suppression.

It is our aim to design statistical methodology for complex-valued HARDI (\mathbb{C} -HARDI) data that recognize voxels where the diffusion process does not

appear to be symmetric. Such information may be used directly as an imaging biomarker for disease progression or treatment response (Whitcher and Matthews, 2006), or as an input to fiber-tracking algorithms (Lazar, 2010; Chung et al., 2011). We also note that analyzing complex-valued MRI measurements has recently been proposed in the functional MRI (fMRI) literature. Several papers by Rowe and colleagues have proposed to retrieve biological information from the phase of fMRI data (Rowe and Logan, 2004; Rowe, 2005; Rowe et al., 2007). The phase structure in fMRI may be due to vascular effects (Menon, 2002) or neuronal effects (Bodurka et al., 1999).

Including higher-order terms from the Kramer-Moyal expansion of the three-dimensional Boltzmann equation implies that even and odd powers of q [Has q been define?] feature in the expansion. It has been shown that any asymmetry in the diffusion PDF leads to complex-valued observations in q -space, or equivalently odd terms in the Kramer-Moyal expansion (Liu et al., 2005a). Estimation of such terms requires (\mathbb{C} -HARDI) measurements. Acquiring observations at a single non-zero magnitude of b , over the entire sphere (i.e., a single shell), one may estimate only a single linear term in the expansion, just as estimating higher-order even-numbered moments requires several shells. This does not mean we believe in net transport at a given voxel, but the higher-order asymmetric effects in the three-dimensional Boltzmann equation may be approximated by a linear term, over a range of q -values. Determining asymmetry is important for better characterisation of local structure. The ability to distinguish forking, or bending, versus crossing fibers may depend on the use of complex-valued observations (Liu et al., 2005a).

Before interpreting the observed phase in DW-MRI we must model and understand phase contributions that will be observed at the voxel level and on given receiver channel (i.e., a receiver coil and its associated electronics), although the underlying causes will be different from the fMRI case. For simplicity, we refer to all elements involved in receiving the signal as a “coil.” We assume that multiple complex-valued measurements are made at each coil and we aim to estimate the phase and amplitude of the diffusion PDF from such measurements. Following Aksoy et al. (2008) we model movement-induced phase as affecting the smooth components of the data in a given axial slice (x - y plane), permitting us to remove such effects. Remaining phase perturbations are either attributed to an imperfect procedure for removing the movement phase or susceptibility effects, as well as eddy-current effects, or genuine phase structure. Phase effects are both subject dependent

susceptibility effects, that can be determined from the $b = 0$ images, as well as eddy current induced inhomogeneities that may be detected when using diffusion-weighted ($b \neq 0$) measurements. The eddy currents are not subject dependent (Truong et al., 2008). Once the bias due to movement has been removed from each coil, a weighted average is formed to estimate the diffusion PDF phase in q -space, and a simple sum is subsequently used to estimate the amplitude of the diffusion PDF. This yields a single amplitude and a single phase for each voxel and diffusion encoding gradient direction.

If there are no susceptibility effects to remove, then the phase must observe constraints from the Hermitian symmetry of the data. The phase at $b = b'$, the nonzero b -value over which q -space shell measurements have been collected, should have a zero “shell average.” We may test whether this null hypothesis may be rejected. If the zero average of the phase is not rejected at $b > 0$, then we assume that the raw phase has been appropriately processed to remove artifacts. We estimate the coefficient of a Taylor expansion of the three-dimensional Boltzmann equation, and test for significant Hermitian structure. If this is the case, then we may estimate the Hermitian structure and try to determine what additional information this may give us regarding the asymmetry of the diffusion PDF.

We discuss theoretical PDFs which will lead to Hermitian (complex) structure in q -space and introduce analytical models for bending and forking fibers. We show voxels identified in \mathbb{C} -HARDI data, acquired using clinical settings, that possess Hermitian structure have a similar appearance in q -space to theoretical models of diffusion PDFs, and discuss to what spatial structure this would correspond. Given our extensive simulation studies, as well as analysis of clinical data, we believe there is clear evidence of the utility in the phase information from \mathbb{C} -HARDI data.

2. Theory

2.1. Modelling complex-valued HARDI observations

At any location \mathbf{x}_n , gradient \mathbf{q}_l we define the q -space observation

$$Z(\mathbf{q}_l; \mathbf{x}_n) = \mathcal{Z}(\mathbf{q}_l; \mathbf{x}_n) + \epsilon(\mathbf{q}_l; \mathbf{x}_n) \equiv A(\mathbf{q}_l; \mathbf{x}_n) \exp(-i\phi(\mathbf{q}_l; \mathbf{x}_n)), \quad (1)$$

$n = 1, \dots, N^2$, $l = 1, \dots, L$, where $\mathcal{Z}(\mathbf{q}_l; \mathbf{x}_n)$ is the noise free (complex-valued) diffusion process and $\epsilon(\mathbf{q}_l; \mathbf{x}_n)$ is complex proper¹ Gaussian noise with variance $\sigma^2(\mathbf{x}_n)$, uncorrelated in both indices l and n , and constant in magnitude over the directions l (Picinbono, 1996). We assume that for $l = 1, \dots, L_0$, the observations are collected at $b = 0$, and for $l = L_0+1, \dots, L$, the observations are collected at $b = b' > 0$. To facilitate exposition we introduce the polar decomposition

$$\mathcal{Z}(\mathbf{q}_l) = \mathcal{A}(\mathbf{q}_l) \exp(-i\varphi(\mathbf{q}_l)) \quad (2)$$

of the noise-free observations, where $\mathcal{A}(\mathbf{q}_l; \mathbf{x}_n)$ and $\varphi(\mathbf{q}_l; \mathbf{x}_n)$ denote the amplitude and phase terms, respectively. We need to model the structure of this object in terms of gradient directions l . The process $\mathcal{Z}(\mathbf{q}_l; \mathbf{x}_n)$ is the Fourier transform (FT) of the spatial PDF (probability density function) at location \mathbf{x}_n , given by

$$a(\mathbf{x}) = \int_{\mathbb{R}^3} \mathcal{Z}(\mathbf{q}; \mathbf{x}_n) \exp(2i\pi \mathbf{x}^T \mathbf{q}) d^3 \mathbf{q}, \quad (3)$$

where we have dropped the explicit dependence upon n . We may, in general, write

$$a(\mathbf{x}) = a_e(\mathbf{x}) + a_o(\mathbf{x}), \quad (4)$$

at a fixed voxel \mathbf{x}_n , where $a_e(\mathbf{x})$ is an *even* function in \mathbf{x} , its FT being real-valued, and $a_o(\mathbf{x})$ is an *odd* function in \mathbf{x} , its FT being purely imaginary when non-zero. One example of such a scenario would be a branching fiber, where the even function corresponds to the fiber continuing through the voxel, and the odd function corresponds to the new branch of the fiber. Calculating the FT of Eq. (4) gives

$$\mathcal{Z}(\mathbf{q}) = \mathcal{A}_e(\mathbf{q}) - i\mathcal{A}_o(\mathbf{q}) = \mathcal{A}(\mathbf{q}) \exp(-i\varphi(\mathbf{q})), \quad (5)$$

where both $\mathcal{A}_e(\mathbf{q})$ and $\mathcal{A}_o(\mathbf{q})$ are real-valued quantities; i.e., the FTs of $a_e(\mathbf{x})$ and $ia_o(\mathbf{x})$, respectively. Thus, the asymmetry information is encoded in the imaginary part of the q -space measurements $\mathcal{Z}(\mathbf{q})$.

In order to proceed we require a model-based procedure to estimate the phase of the diffusion. Starting from the three-dimensional Boltzmann equation we may expand the observed signal using cumulant expansions and the

¹This comment is strictly speaking not true. Since we would like to reduce acquisition time, the full pass-band is not sampled equally. This makes the noise structure more complicated, but we may ignore this slight inaccuracy.

Kramer-Moyal formula

$$\frac{\mathcal{Z}(\mathbf{q})}{\mathcal{Z}(\mathbf{0})} = \exp \left[-iP^{(\text{im})}(\mathbf{q}) + P^{(\text{re})}(\mathbf{q}) + \dots \right]. \quad (6)$$

(Callaghan, 1993; Liu et al., 2004, 2005a). Since we are modelling the phase and amplitude it is convenient to use the two functions

$$P^{(\text{re})}(\mathbf{q}) = -\frac{1}{2!} \sum_{l_1, l_2} D_{l_1, l_2}^{(2)} q_{l_1} q_{l_2} + \dots \quad (7)$$

$$P^{(\text{im})}(\mathbf{q}) = \sum_{l_1} D_{l_1}^{(1)} q_{l_1} - \frac{1}{3!} \sum_{l_1} \sum_{l_2} \sum_{l_3} D_{l_1, l_2, l_3}^{(3)} q_{l_1} q_{l_2} q_{l_3} + \dots, \quad (8)$$

where $D_{\bullet}^{(n)}$ is the n th order cumulant of \mathbf{x} , $\mathbf{q} = (q_{l_1}, q_{l_2}, q_{l_3}) = \gamma \delta G$ with G being the applied diffusion gradient, γ the gyromagnetic ratio and δ the duration of the two gradient pulses used. The terms $P^{(\text{re})}(\mathbf{q})$ and $P^{(\text{im})}(\mathbf{q})$, in general, contain an infinite number of terms and cannot be estimated from finite sample sizes of C-HARDI data. To find simple expressions over a range of \mathbf{q} values we approximate the functions using

$$P^{(\text{re})}(\mathbf{q}) = -\frac{1}{2} \mathbf{q}^T \mathbf{D}^{(\text{re})} \mathbf{q} + e^{(\text{re})}(\mathbf{q}), \quad (9)$$

$$P^{(\text{im})}(\mathbf{q}) = P_q^{(\text{im})}(\mathbf{q}) \quad \text{if} \quad \|\mathbf{q}\| = q, \quad (10)$$

$$P_q^{(\text{im})}(\mathbf{q}) = \Phi_0^{(q)} + \mathbf{q}^T \mathbf{d}^{(\text{im})} + e_q^{(\text{im})}(\mathbf{q}), \quad (11)$$

where $\Phi_0^{(q)}$ is the phase-shift term and may pick up the remaining bias in the data that the pre-processing steps have not resolved (Section 3.1). The terms

$$\mathbf{d}^{(\text{im})} = \left(D_1^{(1)} \quad D_2^{(1)} \quad D_3^{(1)} \right)^T, \quad (12)$$

$$\mathbf{D}^{(\text{re})} = \begin{pmatrix} D_{1,1}^{(2)} & D_{1,2}^{(2)} & D_{1,3}^{(2)} \\ D_{2,1}^{(2)} & D_{2,2}^{(2)} & D_{2,3}^{(2)} \\ D_{3,1}^{(2)} & D_{3,2}^{(2)} & D_{3,3}^{(2)} \end{pmatrix}, \quad (13)$$

provide convenient matrix notation for the first- and second-order cumulants. The function $e^{(\bullet)}(\mathbf{q})$ is the difference between the theoretical model and our simple approximation, and so

$$e^{(\text{im})}(\mathbf{q}) = P_q^{(\text{im})}(\mathbf{q}) - \Phi_0^{(q)} + \mathbf{q}^T \mathbf{d}^{(\text{im})} \quad (14)$$

$$e^{(\text{re})}(\mathbf{q}) = P^{(\text{re})}(\mathbf{q}) + \frac{1}{2} \mathbf{q}^T \mathbf{D}^{(\text{re})} \mathbf{q}. \quad (15)$$

Because the noise is added to $\mathcal{Z}(\mathbf{q})$ and the truncation errors occur in $\log \mathcal{Z}(\mathbf{q})$ it is not necessarily advantageous to combine them together. Thus, we retain a simple quadratic function for the even terms (i.e., use the standard diffusion tensor model for the even terms), and use a linear approximation for the odd terms, suitable when $\|\mathbf{q}\| = q$.

SOFIA SAYS: This is inconsistent with our previous notation, q was q' ?

If we collect data at several values of q , the diffusion tensor model for the magnitude may need augmentation with higher-order even terms (as in diffusion kurtosis imaging). The function $P^{(\cdot)}(\mathbf{q})$ only admits a valid interpretation over a limited range of q -values. For example, we do not believe in net transport, but rather assume that higher-order terms are locally approximated by the linear term. With these equations we note that

$$\log \mathcal{Z}(\mathbf{q}) = \log \mathcal{A}_0 - i\phi_0 - i\mathbf{q}^T \mathbf{d}^{(\text{im})} - \frac{1}{2}\mathbf{q}^T \mathbf{d}^{(\text{re})} \mathbf{q} + e(\mathbf{q}) \quad (16)$$

$$= \log \mathcal{A}(\mathbf{q}) - i\varphi(\mathbf{q}), \quad (17)$$

where the latter equation is simply the logarithm of Eq. (5), and $\mathcal{A}(\mathbf{q})$ as well as $\phi(\mathbf{q})$ are the noisy (sampled) versions of $\mathcal{A}(\mathbf{q})$ and $\varphi(\mathbf{q})$, so that $\log \mathcal{Z}(\mathbf{q}) = \log \mathcal{A}(\mathbf{q}) - i\phi(\mathbf{q})$. This provides a model for the observations and may be used to construct an estimation procedure. If the complex-valued measurements are real-valued, then there are no asymmetries in the PDF and $\mathbf{d}^{(\text{im})} \equiv \mathbf{0}$. That is, there is nothing of interest in the apparent complexity of the observations and the likelihood may be marginalized in terms of the positive, real-valued random variables $\{A(\mathbf{q})\}$, or we may possibly reduce noise by fitting the parametric model to $\{A_e(\mathbf{q})\}$ rather than to $\{A(\mathbf{q})\}$.

2.2. A bending fiber population model

We now consider a model for the asymmetry of the diffusion PDF at a given voxel; specifically, a voxel that corresponds to a bending structure. In this case it makes sense to think of a single fiber population, which is skewed (bending in space). Unfortunately, to connect this type of structure to the analysis of the magnitude we have to model it as a mixture of two fiber populations and write for

$$\frac{\mathcal{Z}(\mathbf{q})}{\mathcal{Z}(\mathbf{0})} = \alpha \frac{\mathcal{Z}_1(\mathbf{q})}{\mathcal{Z}_1(\mathbf{0})} + (1 - \alpha) \frac{\mathcal{Z}_2(\mathbf{q})}{\mathcal{Z}_2(\mathbf{0})}, \quad (18)$$

where $\alpha = \mathcal{Z}_1(\mathbf{0})/[\mathcal{Z}_1(\mathbf{0}) + \mathcal{Z}_2(\mathbf{0})] \in [0, 1]$, is the volume fraction of each branch (normally would be modelled as a half). Both $\mathcal{Z}_1(\mathbf{q})$ and $\mathcal{Z}_2(\mathbf{q})$ may be modelled as branches of fibers (i.e., highly asymmetric diffusion PDFs) meeting at the center of the voxel. Therefore, they are genuinely complex-valued quantities since each branch is highly asymmetric and is modelled as entering but not exiting the voxel. We then have

$$\varphi(\mathbf{q}) = -\tan^{-1} \left(\frac{\alpha \Im \{ \mathcal{Z}_1(\mathbf{q}) / \mathcal{Z}_1(\mathbf{0}) \} + (1 - \alpha) \Im \{ \mathcal{Z}_2(\mathbf{q}) / \mathcal{Z}_2(\mathbf{0}) \}}{\alpha \Re \{ \mathcal{Z}_1(\mathbf{q}) / \mathcal{Z}_1(\mathbf{0}) \} + (1 - \alpha) \Re \{ \mathcal{Z}_2(\mathbf{q}) / \mathcal{Z}_2(\mathbf{0}) \}} \right) \quad (19)$$

This equation represents the phase structure of the entire voxel of the characteristic function in terms of the volume fraction and the characteristic function of each branch. The phase of the normalized mixture reflects the structure of both branches, depending on the value of α and the value of $|\mathcal{Z}_l(\mathbf{q})|$, $l = 1, 2$. Note that for most of \mathbf{q} -space only one of the two branches is non-zero. The phase from each of the branches is changing perpendicularly to the great circle associated with the magnitude of the branch. The mixture of fiber populations have a phase which is varying parallel to the amplitude.

To exemplify typical bending structure we first construct a single branch by dividing a Gaussian PDF into two halves, non-zero for $x_1 > 0$ and zero for $x_1 < 0$ (without loss of generality). This corresponds to the tensor product of a half-Gaussian (or half-Normal) distribution in x_1 and two Gaussian distributions in x_2 and x_3 , respectively. We model a bending fiber by aggregating two half-Gaussian functions with different axes of the asymmetry. If we used the same axes and eigenvalues we would just get back a single DTI model without bending. We plot the real-part, imaginary-part, amplitude and phase for our example and a choice of parameters in Fig. 1. Observe that the phase in Figs. 1a and 1e is nearly linear, the magnitude reflects the two-tensor structure of the data (or equally a scalene diffusion may be fitted from a single shell of observations), and the magnitude of the imaginary component has similar structure to a DTI model but with two large eigenvalues and one small eigenvalue. **[More text is needed to describe the figure!!!]** The magnitude of the real part is also consistent with a two-tensor model or equally a scalene diffusion for single shell observations.

The bending fiber (in general) has to show variation in the phase mainly perpendicular to the great circle, with its zeros going through the direction of the averages of the two populations. Furthermore, if we are in a voxel where $|\mathcal{Z}_1(\mathbf{q})| \gg |\mathcal{Z}_2(\mathbf{q})|$, then **[I have mucked around with this equation...**

please double check!]

$$\varphi(\mathbf{q}) = -\tan^{-1} \left(\frac{\alpha |\mathcal{Z}_1(\mathbf{q})|^{-1} \Im\{\mathcal{Z}_1(\mathbf{q})\} + (1-\alpha) |\mathcal{Z}_2(\mathbf{q})|^{-1} \Im\{\mathcal{Z}_2(\mathbf{q})\}}{\alpha |\mathcal{Z}_1(\mathbf{q})|^{-1} \Re\{\mathcal{Z}_1(\mathbf{q})\} + (1-\alpha) |\mathcal{Z}_2(\mathbf{q})|^{-1} \Re\{\mathcal{Z}_2(\mathbf{q})\}} \right) \approx \varphi_1(\mathbf{q}), \quad (20)$$

namely the phase of $\mathcal{Z}_1(\mathbf{q})$. To elaborate on this, if (as in the previous example with the half-Gaussian) we constructed $\mathcal{Z}_1(\mathbf{q})$ by taking a symmetric diffusion PDF and setting half of the PDF to zero in the direction \mathbf{v}_1 , then $a_1(\mathbf{x}) = 2a_{1e}(\mathbf{x})I(\mathbf{x}^T \mathbf{v}_1 > 0)$, where $a_{1e}(\mathbf{x})$ has FT $\mathcal{A}_{1e}(\mathbf{q})$. We assume that $\mathcal{A}_{1e}(\mathbf{q})$ is an ellipsoid density with axes $\{\mathbf{v}_k\}$ for $k = 1, 2, 3$, while $\mathcal{A}_{2e}(\mathbf{q})$ is an ellipsoid density with axes $\{\boldsymbol{\xi}_k\}$ for $k = 1, 2, 3$. Functions that are zero for a given half-space are related to the partial Hilbert transform (Granlund and Knutsson, 1995) calculated in some set direction, and we denote the Hilbert transform in direction \mathbf{v} by $\mathcal{H}_{\mathbf{v}}\{\cdot\}$. We define this transform (i.e., the partial Hilbert Transform in direction \mathbf{v}) in space via

$$\mathcal{F}^{-1} \{ \mathcal{H}_{\mathbf{v}} \{ \mathcal{A} \} \} (\mathbf{x}) \equiv -i \operatorname{sign}(\mathbf{x}^T \mathbf{v}) a_{1e}(\mathbf{x}). \quad (21)$$

We can therefore write the characteristic function of an asymmetric component of the bending fiber PDF as

$$\begin{aligned} \mathcal{A}_1(\mathbf{q}) &= \mathcal{A}_{1e}(\mathbf{q}) - i \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{q}) \\ &= \mathcal{A}_{1e}(\mathbf{0}) + \nabla \mathcal{A}_{1e}(\mathbf{0})^T \mathbf{q} + \dots - i (0 + \nabla \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0})^T \mathbf{q} + \dots) \\ &= 1 + \dots - i \left[(\nabla \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}))^T \mathbf{q} + \dots \right], \end{aligned} \quad (22)$$

and so $\varphi_1(\mathbf{q}) \approx (\nabla \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}))^T \mathbf{q}$. We assumed that $a_{1e}(\mathbf{x})$ was even in \mathbf{x} and so its Fourier transform is even and has a maximum at $\mathbf{q} = \mathbf{0}$. This implies that $(\nabla \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}))^T$ is proportional to \mathbf{v}_1 and with $\nabla_{\mathbf{v}}$ denoting the directional derivative in direction \mathbf{v} , we may write down the following relations

$$\nabla_{\mathbf{v}_1} \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}) = C_1 \neq 0 \quad (23)$$

$$\nabla_{\mathbf{v}_2} \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}) = \nabla_{\mathbf{v}_3} \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}) = 0. \quad (24)$$

If $|\mathcal{Z}_1(\mathbf{q})| \approx |\mathcal{Z}_2(\mathbf{q})|$ then instead we have

$$\begin{aligned} \mathcal{A}(\mathbf{q}) &\approx 1 + \dots - i \left[\alpha (\nabla \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}))^T \mathbf{q} + (1-\alpha) (\nabla \mathcal{H}_{\boldsymbol{\xi}_1} \mathcal{A}_{2e}(\mathbf{0}))^T \mathbf{q} + \dots \right] \\ &\approx 1 + \dots - i \left[(\alpha \nabla \mathcal{H}_{\mathbf{v}_1} \mathcal{A}_{1e}(\mathbf{0}) + (1-\alpha) \nabla \mathcal{H}_{\boldsymbol{\xi}_1} \mathcal{A}_{2e}(\mathbf{0}))^T \mathbf{q} + \dots \right], \end{aligned} \quad (25)$$

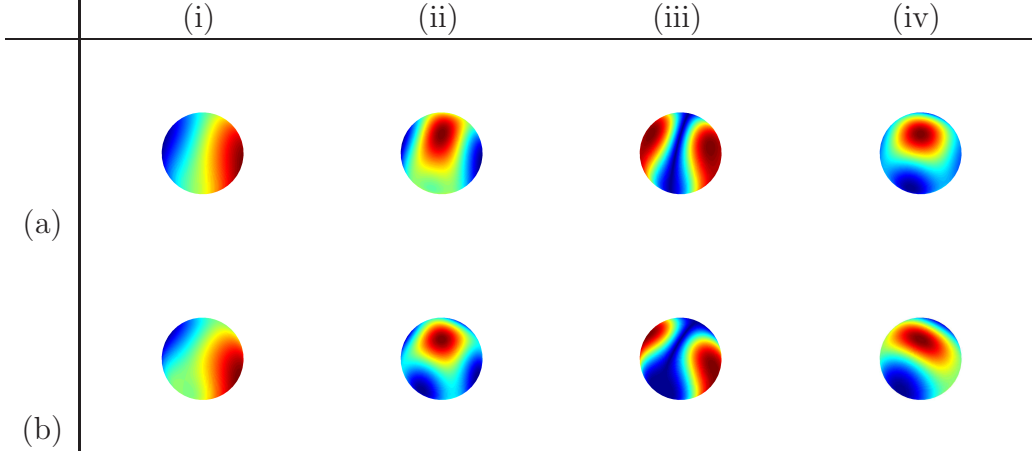


Figure 1: Representations of a simulated bending fiber population for a single voxel in q -space. The rows correspond to (a) a 60° bend and (b) a 90° bend. The columns correspond to (i) the phase, (ii) the absolute value of the real part of the complex-valued signal, (iii) the absolute value of the imaginary part of the complex-valued signal and (iv) the amplitude squared of the complex-valued signal. The phase displays typical asymmetry about the great circle (**defined how?**).

and the total phase is given by

$$\varphi(\mathbf{q}) \approx [\alpha C_1 \mathbf{v}_1 + (1 - \alpha) C_2 \boldsymbol{\xi}_1]^T \mathbf{q}. \quad (26)$$

The average gradient vector is the sum of the bending orientations, and indicates the orientation of the bending. A typical model might be $\alpha = 1/2$, $C_1 = C_2$ and we show an example of this in Fig. 2, where the mean of the vectors show the net apparent shift over a range of q -values. In this case

$$\varphi(\mathbf{q}) \approx \frac{1}{2} C (\mathbf{v}_1 + \boldsymbol{\xi}_1)^T \mathbf{q}, \quad (27)$$

where the orientation of $(\mathbf{v}_1 + \boldsymbol{\xi}_1)$ is the direction of the bending (Fig. 2). The magnitude $C\|\mathbf{v}_1 + \boldsymbol{\xi}_1\|/2$ is proportional to the degree of asymmetry of the diffusion PDF. We reiterate that $\varphi(\mathbf{q})$ is not a mean shift over all scales but a local approximation to the asymmetric structure, over a single shell in q -space.

2.3. A forking fiber population model

We consider another model for the asymmetry of the diffusion PDF at the voxel level that corresponds with a forking structure. In this case we again

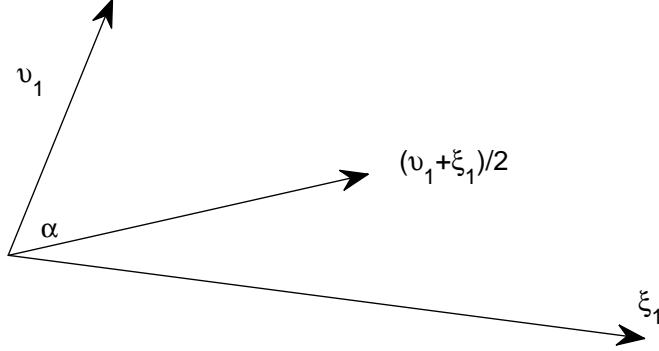


Figure 2: Illustration of the vector of phase. We assume that the bending fiber has two principal orientations, \mathbf{v}_1 and $\boldsymbol{\xi}_1$. If the volume fractions are equal and the derivatives of the PDFs at frequency zero are equal, then the phase vector will be proportional to the average of these two fibers. Hence, the phase vector will point in the direction of bend $(\mathbf{v}_1 + \boldsymbol{\xi}_1)/2$.

need to adjust Eq. (6) to that of a multi-population Eq. (18), where $Z_1(\mathbf{q})$ is the fiber population that runs through the entire voxel and $Z_2(\mathbf{q})$ is the branching fiber population. In this scenario $\mathcal{Z}_1(\mathbf{q}) \in \mathbb{R}$, as this population is symmetric (unless we wish the fork to have two equal branches, in which case the forking would be modelled as aggregating three asymmetric branches). The phase of the observed signal intensities is given by

$$\varphi(\mathbf{q}) = -\tan^{-1} \left(\frac{\Im\{\mathcal{Z}(\mathbf{q})\}}{\Re\{\mathcal{Z}(\mathbf{q})\}} \right) \quad (28)$$

$$= -\tan^{-1} \left(\frac{(1 - \alpha)\Im\{\mathcal{Z}_2(\mathbf{q})\}}{\alpha\mathcal{Z}_1(\mathbf{q}) + (1 - \alpha)\Re\{\mathcal{Z}_2(\mathbf{q})\}} \right). \quad (29)$$

The concentration of $Z_1(\mathbf{q})$ and $Z_2(\mathbf{q})$ will not be uniform on the sphere in q -space. The most reliable structure in the imaginary component of the diffusion PDF will occur when $|Z_2(\mathbf{q})| \gg |Z_1(\mathbf{q})|$. In this case the phase is changing perpendicularly to the great circle associated with $Z_2(\mathbf{q})$, but there is no phase associated with $Z_1(\mathbf{q})$. We plot an example of this in Fig. 3. We observe that the phase is more complicated than in the bending case, and cannot be approximated as simply linear. The amplitude is still that of a mixture of two PDFs, while the magnitude of the imaginary part is like a DTI model with two large eigenvalues and one small. The real part is similar to the magnitude.

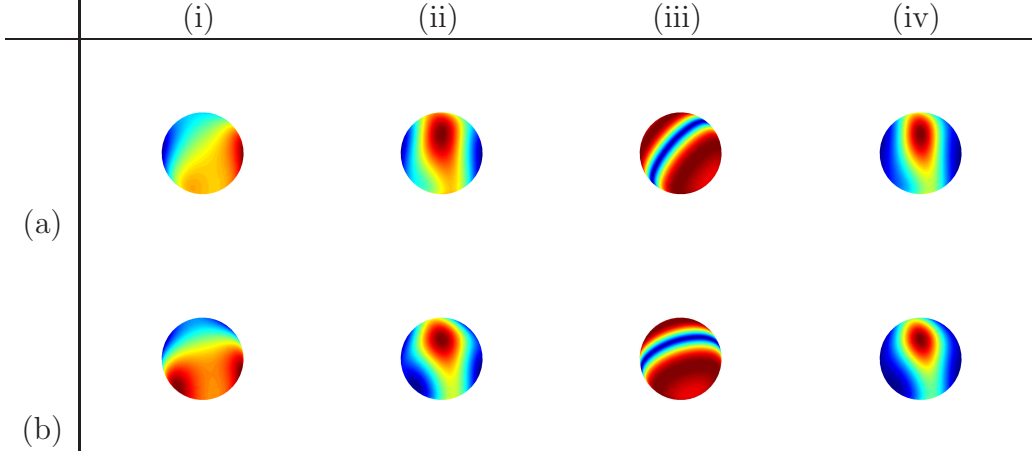


Figure 3: Representations of a simulated forking fiber population for a single voxel in q -space. The rows correspond to (a) a 60° bend and (b) a 90° bend. The columns correspond to (i) the phase, (ii) the absolute value of the real part of the complex-valued signal, (iii) the absolute value of the imaginary part of the complex-valued signal and (iv) the amplitude squared of the complex-valued signal.

Utilizing the same Taylor series expansion as in the bending case, we have

$$\mathcal{A}(\mathbf{q}) \approx 1 + \dots + i \left[(1 - \alpha) (\nabla \mathcal{H}_{\nu_2} \mathcal{A}_{2e}(\mathbf{0}))^T \mathbf{q} + \dots \right],$$

and the total phase is given by

$$\varphi(\mathbf{q}) \approx (1 - \alpha) (\nabla \mathcal{H}_{\xi_1} \mathcal{A}_{2e}(\mathbf{0}))^T \mathbf{q} \quad (30)$$

$$= (1 - \alpha) C_2 \xi_1^T \mathbf{q} \quad (31)$$

The orientation of the forking fiber is therefore determined by the direction of $\varphi(\mathbf{q})$, or what direction the new fork is going into, and the magnitude given by the volume fraction and the derivative of the phase at zero wavenumber. The sign of the vector tells us the parity of the forking in a given orientation and provides additional orientational information.

3. Methods

3.1. Pre-processing \mathbb{C} -HARDI data

In practice the observations do not take the form of Eq. (1). Instead of a single measurement at each spatial location and gradient orientation,

we have measurements from a number of coils, that shall here be denoted by $\gamma = 1, \dots, N_\gamma$ (recall that by coil we actually refer to receiver coils and their associated electronics). These observations are all complex-valued, but have varying sensitivity over space. In practice, measurements from all coils must be combined to get an acceptable SNR at the voxel level in the brain. All the different coils are subjected to a random phase-shift common to all the spatial locations, and when treating the volumes slice-by-slice for fixed superior-inferior coordinates, then a common movement induced phase may also be modelled.

The nature of phase must be modelled in order to separate contributions to the q -space observations mainly from inter-scan or intra-scan motion (Aksoy et al., 2008)), susceptibility effects and true excess phase due to the observation of asymmetric diffusion PDFs. Some discussion of motion-induced artifacts causing phase variation may also be found in; e.g., Liu et al. (2005b), Bretthorst (2006) and Aksoy et al. (2008).

We model the observations at each coil and gradient as a multiplicative relationship between the true amplitude, a coil sensitivity effect and an observed phase effect. These observations are assumed to have a fixed variance over gradient directions at each voxel. We argue that phase due to motion is a large-scale phenomenon, and estimate this source phase variability using the low-frequency portion of phase from each coil. **We shall discuss the susceptibility effects further in section X.** This yields an estimate of the phase at each coil and gradient that needs to be combined to an estimate at each gradient, based on the stochastic properties of the raw estimates. A weighted average is performed across the coil phases at each gradient, by weighting each phase by the magnitude of the observation at that coil and gradient combination. The details for this entire process is provided in Appendix A. We also aggregate the magnitudes over the coils to estimate a single magnitude at each gradient direction. This procedure yields an estimator of the complex-valued observations at each gradient orientation and spatial location whose properties can be approximated as corresponding to Eq. (1).

3.2. Least squares estimation of phase

Three important questions are addressed in this section:

1. Is there remaining (non-Hermitian) mean phase structure at $\|\mathbf{q}\| = q$ which cannot be explained by known (artifact-based) causes? This is a validation of the pre-processing steps.

2. Is there remaining Hermitian phase structure at $\|\mathbf{q}\| = q$ which is statistically significant?
3. Can we explain the Hermitian structure using a parsimonious model, and can we estimate the orientation of the asymmetry?

In order to address these questions we must estimate the structure of the observed \mathbb{C} -HARDI data. With a single estimated amplitude and phase at each voxel, the parameters of the structure in Eq. (16) may be estimated. The mean of the real and imaginary part may be estimated separately, forming the matrices

$$\begin{aligned}
\mathbf{A} &= (\log |Z(\mathbf{q}_1)| \quad \cdots \quad \log |Z(\mathbf{q}_L)|)^T, \\
\mathbf{\Phi} &= (\phi(\mathbf{q}_{L_0+1}) \quad \cdots \quad \phi(\mathbf{q}_L))^T, \\
\mathbf{P} &= \begin{pmatrix} 1 & -bq_{11}^2 & -bq_{12}^2 & \cdots & -bq_{12}q_{13} \\ 1 & -bq_{21}^2 & -bq_{22}^2 & \cdots & -bq_{22}q_{23} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -bq_{L1}^2 & -bq_{L2}^2 & \cdots & -bq_{L2}q_{L3} \end{pmatrix}, \\
\mathbf{R} &= \begin{pmatrix} 1 & q_{L_0+11} & q_{L_0+12} & q_{L_0+13} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & q_{L1} & q_{L2} & q_{L3} \end{pmatrix}, \\
\boldsymbol{\alpha} &= \left(A_0 \quad D_{11}^{(2)} \quad D_{22}^{(2)} \quad D_{33}^{(2)} \quad \cdots \quad D_{23}^{(2)} \right)^T, \\
\boldsymbol{\beta} &= \left(\Phi_0^{(q)} \quad D_1^{(1)} \quad D_2^{(1)} \quad D_3^{(1)} \right)^T.
\end{aligned} \tag{32}$$

$$\boldsymbol{\beta} = \left(\Phi_0^{(q)} \quad D_1^{(1)} \quad D_2^{(1)} \quad D_3^{(1)} \right)^T. \tag{33}$$

We then estimate the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ using the least-squares estimators

$$\hat{\boldsymbol{\alpha}} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{A}, \tag{34}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{\Phi}, \tag{35}$$

respectively. We denote the covariance matrix of \mathbf{A} by $\boldsymbol{\Sigma}_{\mathbf{A}}$ and the covariance matrix of $\mathbf{\Phi}$ as $\boldsymbol{\Sigma}_{\mathbf{\Phi}}$. There are no cross correlations between $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$, as the noise is approximately complex proper. The matrices $\boldsymbol{\Sigma}_{\mathbf{A}}$ and $\boldsymbol{\Sigma}_{\mathbf{\Phi}}$ are both diagonal with the diagonal entries given by

$$\begin{aligned}
\sigma_{kk}^{(A)} = \text{Var} \{ \log |Z(\mathbf{q}_{k+L_0})| \} &= \frac{\sigma^2}{2\mathcal{A}^2(\mathbf{q}_{k+L_0})} \\
&= \text{Var} \{ \phi(\mathbf{q}_{k+L_0}) \} = \sigma_{kk}^{(\Phi)},
\end{aligned}$$

for $k = 1, \dots, L - L_0$. Appealing to the central limit theorem, one can show that $\hat{\alpha}$ and $\hat{\beta}$ converge at a rate of \sqrt{L} , and they are asymptotically normally distributed as $L - L_0 \rightarrow \infty$. This statement is valid unless we implement an acquisition scheme that samples the sphere densely and in a very irregular manner. Following from the theory of least squares we have that

$$\begin{aligned}\Sigma_{\beta} &= \text{Var} \left\{ \hat{\beta} \right\} \\ &= \text{Var} \left\{ (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \Phi \right\} \\ &= (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \Sigma_{\Phi} \mathbf{R} (\mathbf{R}^T \mathbf{R})^{-1}.\end{aligned}\tag{36}$$

We can plug in estimators of all these quantities and then obtain $\hat{\Sigma}_{\beta}$. We have derived estimators for both of the parameters of interest (α and β) and also an estimate for the variance of β .

Note that the DTI model is estimated from the real part of the logarithm of the observations; this is equivalent to performing a regression of the logarithm of the magnitudes, a standard method for estimating the DTI model (Basser et al., 1994b). We also believe the model may not be suitable if the voxel corresponds to cortico-spinal fluid (CSF) or grey matter, as in such cases the distributional assumptions that make least squares methods suitable are not met. We propose to remove such voxels from our analysis based on their fractional anisotropy or mean diffusivity.

3.2.1. Existence of the remaining phase artifacts

We wish to test the hypothesis that the preprocessing was successful; this case corresponds to the average phase over the shell being zero (assuming discretization effects are negligible). We formulate the hypothesis test as

$$H_0 : \Phi_0^{(q)} = 0 \quad \text{versus} \quad H_1 : \Phi_0^{(q)} \neq 0,\tag{37}$$

where $\Phi_0^{(q)}$ is the phase-shift term Eq. (11) from the Kramer-Moyal formula Eq. (6), estimated via linear regression in Eq. (35). Solving the least squares equations explicitly, based on the observed phase terms $\phi(\mathbf{q}_l)$, yields the estimator $\hat{\Phi}_0^{(q)} = (L - L_0)^{-1} \sum_{l=L_0+1}^L \phi(\mathbf{q}_l)$, and the variance of the estimator

is given by

$$\text{Var} \left\{ \widehat{\Phi}_0^{(q)} \right\} = \frac{1}{(L - L_0)^2} \sum_{l=L_0+1}^L \text{Var} \{ \phi(\mathbf{q}_l) \} \quad (38)$$

$$= \frac{1}{(L - L_0)^2} \sum_{l=L_0+1}^L \frac{\sigma^2}{2\mathcal{A}^2(\mathbf{q}_l)}. \quad (39)$$

To estimate the components of this expression we need good estimators of σ^2 and $\mathcal{A}^2(\mathbf{q}_l)$. We use the method of moments estimator of the observed amplitude squared for the latter quantity, and introduce a robust estimator of σ^2 given by median absolute deviation (MAD)

$$\hat{\sigma}_{\text{MAD}} = \sqrt{2} \frac{\text{median} \{ |\phi(\mathbf{q}_l) - \phi(\mathbf{q}_{b(l)})| A(\mathbf{q}_l) \}}{0.6745}, \quad (40)$$

where $b(l)$ is the nearest neighbor of index l . The factor of 0.6745 is standard to construct an unbiased estimator for Gaussian data (Hoaglin et al., 1983). To obtain a valid test statistic we therefore take

$$T_1(q) = \frac{\widehat{\Phi}_0^{(q)} \sqrt{2} A(\mathbf{q}_l)}{\hat{\sigma}_{\text{MAD}}} \stackrel{d|H_0}{=} N(0, 1). \quad (41)$$

We can therefore (at each voxel) test the hypotheses that $\Phi_0^{(q)} = 0$; by calculating an observed value of T_1 and comparing this to the standard Gaussian distribution. If the average phase is not zero, then this is indicative of either insufficient removal of phase induced by movement or remaining susceptibility effects. In either case, the Hermitian structure is not in this case interpretable, as deviating from a non-zero mean. These voxels should be removed from further analysis.

3.2.2. Existence of Hermitian phase structure

The second hypothesis test corresponds to testing whether the Hermitian structure is non-negligible

$$H_0 : \mathbf{d}^{(\text{im})} = \mathbf{0} \quad \text{versus} \quad H_1 : \mathbf{d}^{(\text{im})} \neq \mathbf{0}. \quad (42)$$

Recall that $\mathbf{d}^{(\text{im})}$ is the vector of first-order cumulants Eq. (12) of the Kramer-Moyal formula Eq. (6), estimated via linear regression in Eq. (35). This test

is formulating the question, is there any asymmetry in the diffusion PDF, or is it symmetric? Define $\mathbf{F} = \text{diag}(0, 1, 1, 1)$, and then the Wald statistic for this problem is given by

$$T_2 = \left(\mathbf{F}\hat{\boldsymbol{\beta}}\right)^T \left(\mathbf{F}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}}\mathbf{F}^T\right)^{-1} \mathbf{F}\hat{\boldsymbol{\beta}} \stackrel{d|H_0}{=} \chi_3^2. \quad (43)$$

This statistic provides us with a method to test for Hermitian structure in the phase. The information of the $\mathbf{d}^{(\text{im})}$ vector is important for two reasons: it provides us with a method of detecting Hermitian structure present in the data and we may estimate the orientation of that direction and so both the magnitude and the directionality of this object is of interest.

3.2.3. Existence of linear phase structure

Finally, having established that no artificial or random phase shifts remain, and that Hermitian structure can be found in the data, we may wish to try to estimate this structure by a moderate number of parameters. We saw from our previous arguments that bending fibers have near linear phase, whilst forking/fanning fibers have more complicated phase. We therefore fit a linear term to the phase, to capture the main hermitian structure. If the structure is consistent with bending, then the linear terms should capture most of the structure in the phase. In this case R^2 (the coefficient of determination) from fitting the linear phase model in Eq. (35) should take a value near one. By calculating R^2 we may determine how near linear the phase is.

We define R^2 to be the fraction of energy explained by the linear term

$$R^2 = \frac{\sum_{l=1}^{L-L_0} \left(\hat{d}_1^{(q)} q_{L_0+l,1} + \hat{d}_2^{(q)} q_{L_0+l,2} + \hat{d}_3^{(q)} q_{L_0+l,3} \right)^2}{\sum_{l=1}^{L-L_0} \phi_l^2}. \quad (44)$$

Note that when estimating \mathbf{d} we also estimate the $\Phi_0(q)$ term, and so no non-zero mean phase terms will bias our estimation of \mathbf{d} (it is poor statistical practice to estimate the linear term before the main effect). Recall that we want to keep a potentially non-zero main effect in the model (insufficient phase removal) and then see what fraction of energy is explained by the linear (Hermitian) terms. A Taylor series expansion may justify this (*cf.* Eq. (11)), and see also the plots in Figs. 1 and 3. Since it is true that $\sum_j \mathbf{q}_j \approx \mathbf{0}$ it however makes little difference whether one includes a mean term or not when estimating the linear terms. Note that unlike DTI we must sample the entire sphere instead of a half-sphere to estimate this effect.

3.3. Model comparison

As a real-valued method of comparison we shall fit three possible models, namely the isotropic model of

$$\mathcal{A}(\mathbf{q}_l) = \mathcal{A}(q), \quad l = L_0 + 1, \dots, L, \quad (45)$$

the single tensor model of

$$\mathcal{A}(\mathbf{q}_l) = \exp \left\{ -\frac{1}{2} \mathbf{q}^T \mathbf{D}^{(\text{re})} \mathbf{q} \right\}, \quad (46)$$

$l = 1, \dots, L$, as well as the two tensor model of

$$\mathcal{A}(\mathbf{q}_l) = \frac{1}{2} \exp \left\{ -\frac{1}{2} \mathbf{q}^T \mathbf{D}_1^{(\text{re})} \mathbf{q} \right\} + \frac{1}{2} \exp \left\{ -\frac{1}{2} \mathbf{q}^T \mathbf{D}_2^{(\text{re})} \mathbf{q} \right\}, \quad (47)$$

$l = L_0 + 1, \dots, L$. Note that we have in the latter case enforced equal weighting between fibers of a half. This is to ensure model identifiability. Basically, if we only collect measurements at $b = 0$ and a single shell we can receive the same value for $a_1 \exp\{-\mathbf{q}^T \mathbf{D}_1^{(\text{re})} \mathbf{q}/2\}$ by equivalently changing the scaling of $\mathbf{D}_1^{(\text{re})}$ or the volume fraction. To ensure identifiability we either need to fix the trace of the diffusion matrices or the volume fractions.

For simplicity we first normalize the measurements by the following rule

$$\bar{A}(\mathbf{0}) = \frac{1}{L_0} \sum_{l=1}^{L_0} \sum_{\gamma} A_{\gamma}(\mathbf{q}_l), \quad \tilde{A}(\mathbf{q}_l) = \frac{\sum_{\gamma} A_{\gamma}(\mathbf{q}_l)}{\bar{A}(\mathbf{0})}. \quad (48)$$

This ensures the PDF estimate has been normalized. To produce estimated fitted models we would use the objective function:

$$\hat{\boldsymbol{\theta}} = \arg \min \sum_{l=L_0+1}^L \left[\tilde{A}(\mathbf{q}_l) - \mathcal{A}(\mathbf{q}_l) \right]^2, \quad (49)$$

where $\hat{\boldsymbol{\theta}} = ()$ [WHAT is $\hat{\boldsymbol{\theta}}$] and $\mathcal{A}(\mathbf{q}_l)$ is given by Eqs. (45), (46) or (47). This however does not allow us to easily chose between the models of Eqs. (45), (46) or (47). It is natural that Eq. (47) fits the data better than Eq. (45) because it has more parameters. We denote the number of parameters in any one model by p . We therefore account for “learning of the noise” by using a

model choice procedure. We here apply the Bayesian information criterion (BIC) (Schwarz, 1978), and define the new optimization problem of

$$\left[\hat{\boldsymbol{\theta}}, \hat{p}\right] = \arg \min \left\{ (L - L_0) \log \left[\frac{1}{L - L_0} \sum_{l=L_0+1}^L \left(\tilde{A}(\mathbf{q}_l) - \mathcal{A}(\mathbf{q}_l) \right)^2 \right] + p \log[L - L_0] \right\}. \quad (50)$$

This will permit us to both chose the model and estimate the parameters of the best fitting model. We reiterate that bending fibers are more likely to be seen as single-tensor models, and forking fibers as mixture models.

SOFIA SAYS: We seem to underutilize the simulations. Brandon didn't like the extra plots, so I removed them, but maybe we should use a table or something?

3.4. Clinical data acquisition

In vivo data were acquired on a single healthy male volunteer, who gave informed consent according to local policies, using a GE Signa HDx system (General Electric, Waukshua, WI, USA), with actively shielded magnetic field gradients (maximum amplitude 40 mT/m⁻¹). The body coil was used for RF transmission, and an eight-channel head coil for signal reception. No parallel imaging speedup method was applied in order to simplify data processing. Each volume was acquired using a “peripherally gated” multi-slice singly refocused spin echo EPI sequence, with parameters optimised for precise measurement of the diffusion tensor in parenchyma, with isotropic (2.5×2.5×2.5 mm) voxels and a 128×128 in-plane matrix. Because of scanner limitations, only 19 slices were acquired per volume. Hence, an interslice gap of 2.5 mm was used to ensure most of the brain was surveyed. Repetition time was set to 5 R-R intervals, and the echo time was extended to 150 ms to allow full echo data to be collected and thus avoid the complication of the homodyne reconstruction which would otherwise be necessary.

Based on the recommendations of Jones et al. (1999), the maximum diffusion weighting was 1300 s/mm⁻², and at each slice location, two images were acquired with no diffusion gradients applied, together with 26 diffusion-weighted images in which gradient directions were uniformly distributed in space. As with the number of slice locations, choices for these parameters were severely constrained by scanner limitations. For each volume (i.e., each set of 19 slice locations), complex data was stored as separate real and imaginary images for each of the eight receive coils, along with a coil-combined

(magnitude) image automatically created by the scanner. GE’s “gradwarp” processing (which attempts to correct for minor image distortions due to imaging gradient non-linearities) was disabled for all images.

The coil-combined magnitude images were corrected for the effects of eddy-current induced distortion and subject motion using in-house software. For each data set, a two-stage registration process was used. First, the diffusion-weighted volumes were co-registered together (using the mean diffusion weighted volume as reference) with a 12-parameter 3D affine transformation. The non-diffusion weighted volumes were similarly co-registered, but with a six-parameter rigid body transformation. The co-registered non-diffusion weighted volumes were then registered to the co-registered diffusion-weighted volumes by another six-parameter rigid body transformation. Each registration step used the FLIRT routine, part of FSL (Jenkinson and Smith, 2001; Smith et al., 2004). The transformations determined by this process were then applied to the set of 16 corresponding images from the complex-valued data (i.e., real and imaginary images for each of eight coils).

3.5. Simulations

The simulation of fiber population structures involves the combination of Gaussian and non-Gaussian models as diffusion PDFs. We model this structure directly in q -space. As a starting point, we provide the standard Gaussian diffusion tensor model

$$G(\mathbf{q}; \mathbf{\Lambda}, \mathbf{V}) = \exp \left(-\lambda_1 \langle \mathbf{v}_1, \mathbf{q} \rangle^2 - \lambda_2 \langle \mathbf{v}_2, \mathbf{q} \rangle^2 - \lambda_3 \langle \mathbf{v}_3, \mathbf{q} \rangle^2 \right), \quad (51)$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ is the diagonal matrix of eigenvalues and $\mathbf{V} = \{\mathbf{v}_k\}_k$ is the matrix of eigenvectors. The half-Gaussian diffusion model is given by

$$\begin{aligned} \text{HG}(\mathbf{q}; \mathbf{\Lambda}, \mathbf{V}) &= \exp \left(-\lambda_2 \langle \mathbf{v}_2, \mathbf{q} \rangle^2 \right) \exp \left(-\lambda_3 \langle \mathbf{v}_3, \mathbf{q} \rangle^2 \right) \\ &\times \left[\exp \left(-\lambda_1 \langle \mathbf{v}_1, \mathbf{q} \rangle^2 \right) - i \frac{2}{\sqrt{\pi}} D \left(\sqrt{\lambda_1} \langle \mathbf{v}_1, \mathbf{q} \rangle \right) \right], \end{aligned} \quad (52)$$

where $D(\cdot)$ is the Dawson function (Abramowitz and Stegun, 1972). The phase of the half-Gaussian diffusion model is given by

$$\varphi^{(\text{HG})}(\mathbf{q}) = -\tan^{-1} \left(\frac{-\frac{2}{\sqrt{\pi}} D \left(\sqrt{\lambda_1} \langle \mathbf{v}_1, \mathbf{q} \rangle \right)}{\exp \left(-\lambda_1 \langle \mathbf{v}_1, \mathbf{q} \rangle^2 \right)} \right) \quad (53)$$

$$\approx \frac{2}{\sqrt{\pi}} \sqrt{\lambda_1} \langle \mathbf{v}_1, \mathbf{q} \rangle + \dots \quad (54)$$

We provide a plot of the true phase, and its linear approximation, in Fig. 4 for a number of λ_1 namely $\lambda_1 = 0.54$ (dash-dotted) $\lambda = 2.72$ (solid) and $\lambda = 13.60$ (dotted). The linear phase approximation, represented by the thick line for each line style and λ_1 value, appears to be reasonable for the half-Gaussian diffusion model as long as q does not become too large, in absolute value.

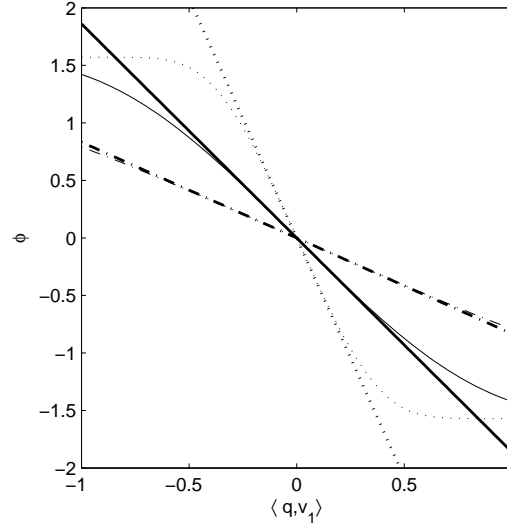


Figure 4: A plot of the phase of the half-Gaussian diffusion model as a function of the principle eigenvalue λ_1 . The approximation given by Eq. (53) is represented as the thick line, for each line style. The value of λ_1 is indicated by the line style, namely $\lambda_1 = 0.54$ (dash-dotted) $\lambda = 2.72$ (solid) and $\lambda = 13.60$ (dotted).

We propose five different fiber population structures in order to investigate the typical characteristics of diffusion PDFs discussed previously. Table 1 provides a summary of these models, where $\mathbf{\Lambda}^{(1)} = 0.04 \cdot \text{diag}(17 \times 4, 8, 8)$, $\mathbf{\Lambda}^{(2)} = 0.04 \cdot \text{diag}(84/3, 84/3, 84/3)$. The orientation of the first diffusion PDFs is chosen randomly, and is given by the basis \mathbf{V} , while the orientation of the second PDF is rotated by an angle of θ in the third axis (q_3).

Table 1: Summary of fiber population structure models used in the simulations. The functions $G(\mathbf{q}_l)$ and $HG(\mathbf{q}_l)$ are defined in Eqs. (51) and (52), respectively.

Name	Notation	Function
Truncated Prolate Diffusion Process	$TPD(\mathbf{q}_l)$	$HG(\mathbf{q}_l; \Lambda^{(1)}, \mathbf{V})$
Prolate Diffusion Process	$PD(\mathbf{q}_l)$	$G(\mathbf{q}_l; \Lambda^{(1)}, \mathbf{V})$
Isotropic Diffusion Process	$ID(\mathbf{q}_l)$	$HG(\mathbf{q}_l; \Lambda^{(2)}, \mathbf{V})$
Forking Diffusion Process	$FD(\mathbf{q}_l)$	$\frac{2}{3}G(\mathbf{q}_l; \Lambda^{(1)}, \mathbf{V}) + \frac{1}{3}HG(\mathbf{q}_l; \Lambda^{(1)}, R_\theta \mathbf{V})$
Bending Diffusion Process	$BD(\mathbf{q}_l)$	$\frac{1}{2}HG(\mathbf{q}_l; \Lambda^{(1)}, \mathbf{V}) + \frac{1}{2}HG(\mathbf{q}_l; \Lambda^{(1)}, R_\theta \mathbf{V})$

4. Results

4.1. Simulated data

We simulate diffusion PDFs following Alexander (2005) in the class of diffusion parameters, but adjust the PDFs to match our choice of $b = 1300$. Both in the simulation studies and the clinical data acquisition we use a set of gradients determined using the method of Jones et al. (1999). The signal-plus-noise model is given by

$$Z(\mathbf{q}_l) = \mathcal{Z}^{(k)}(\mathbf{q}_l) + \sigma\epsilon_l + i\sigma\eta_l \quad (55)$$

$$\mathcal{Z}(\mathbf{q}_l) = \mathcal{A}(\mathbf{q}_l) \exp(-i\varphi^{(k)}(\mathbf{q}_l)), \quad (56)$$

where $\mathcal{Z}(\cdot)$ is one of the fiber population structure models in Section 3.5, ϵ_l and η_l are independent realizations from a Gaussian distribution with equal standard deviations of $TPD(\mathbf{0})/20$. We randomize each trial over $\{\mathbf{v}_1\}$ to avoid consistent directional biases, selecting two perpendicular vectors to this eigenvector and force the error standard deviation to be $\sigma = 1/20$. We also rotate the second basis $\{\boldsymbol{\xi}_j\}_j$ with respect to the first by θ in the q_3 axis, where $\theta \in \{\pi/3, \pi/2\}$. Because \mathbf{v}_1 has a random orientation, no consistent orientational biases are accrued by this procedure.

We simulate the artificial DW-MRI acquisitions 5000 times, and apply parametric methods to calculate summary statistics, as outlined in the previous section. Fig. 5 displays the directions associated with the sparse sampling scheme (26 directions) and dense sampling scheme (133 directions) on the sphere. First, we look at the behavior of the test statistics, the null

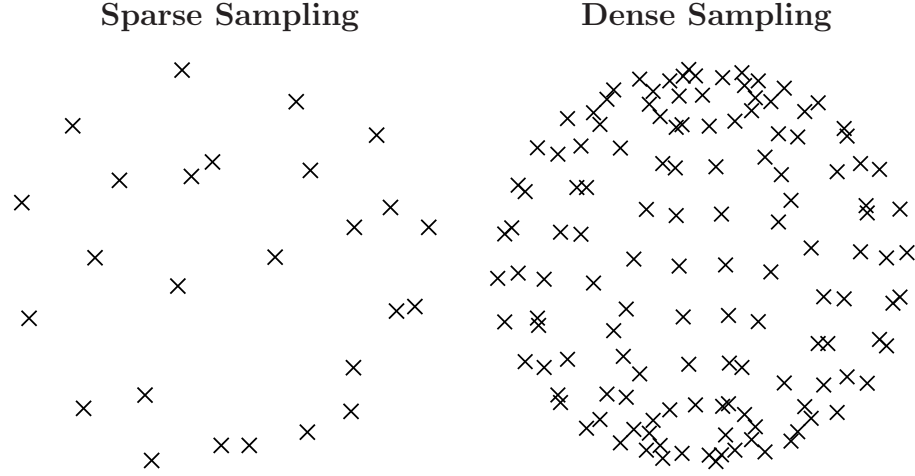


Figure 5: Directional sampling regimes for the simulation study: 26 gradient directions (*sparse sampling*) used in the clinical acquisition (Jones et al., 1999) and a regular sampling of the sphere with 133 gradient directions (*dense sampling*).

average phase T_1 and the Wald statistic T_2 , under the diffusion models in Table 1. When compared with their theoretical distributions, a $N(0, 1)$ for the T_1 statistic and a χ_3^2 for the T_2 statistic, Fig. 6 shows the consistent performance of our statistics. The simulated data should have zero average phase and no Hermitian phase and this is confirmed by the fact that the empirical distributions for each statistic closely match the theoretical distribution. However, we acknowledge the fact that the theoretical distribution for null average phase is more difficult to achieve for certain diffusion models using a sparse sampling of directions over the sphere.

Additional statistics of interest when estimating the linear phase structure are the norm of the imaginary vector $\mathbf{d}^{(\text{im})}$ and the coefficient of determination R^2 . In Fig. 7 it is apparent that the asymmetric diffusion PDFs are clearly distinguished from the symmetric PDFs, by observing different values of the Wald statistic [is this correct?!?] and that the imaginary structure is described by the varying value of the norm of the imaginary vector $\mathbf{d}^{(\text{im})}$. [WHAT?!?]

We would like to distinguish between simple forms of the phase, as would be typical for bending, and more complex forms that cannot be characterised by a linear phase, as in forking, compare with the plots in Figs. 1–3. If one looks at the Figs. 1(a,i) and 1(b,i), and compares them to Figs. 3(a,i)

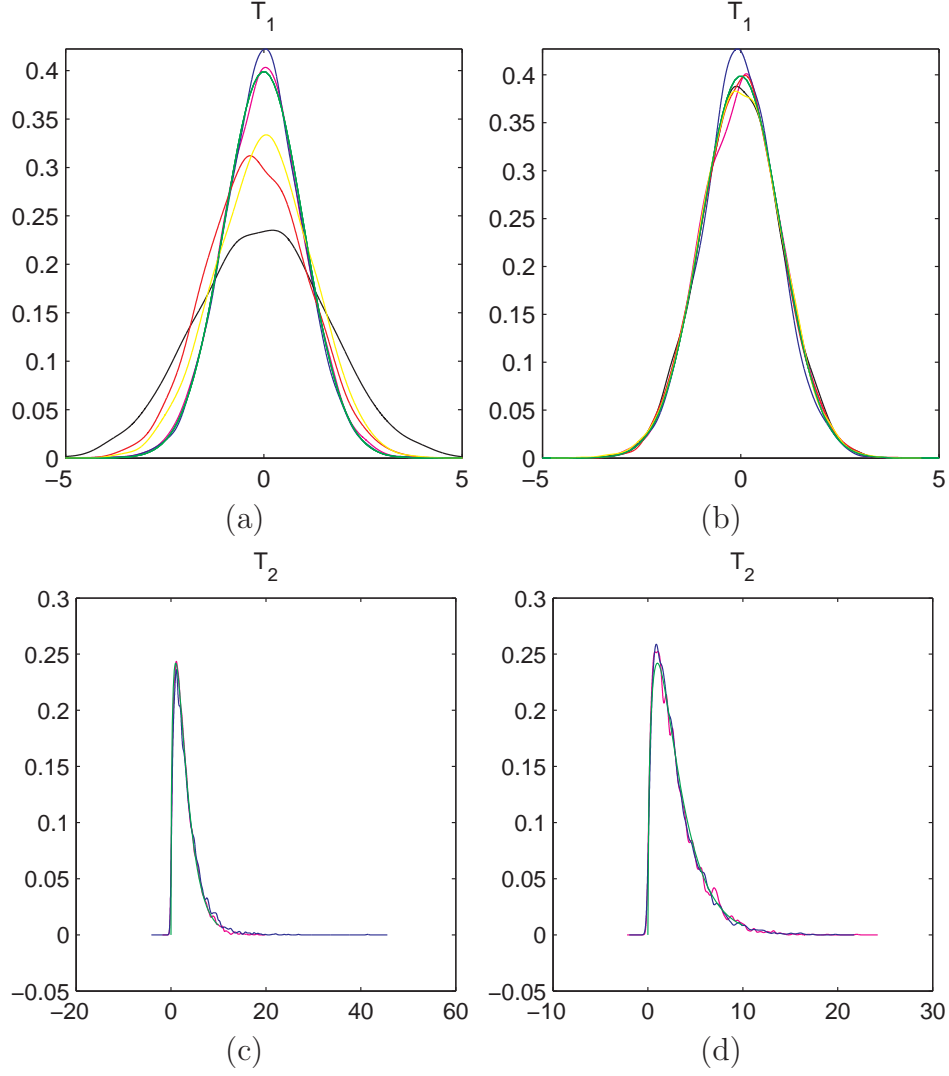


Figure 6: Test statistics T_1 and T_2 from simulations under two q -space sampling schemes. **(a,c)** use the *sparse sampling* scheme and **(b,d)** use the *dense sampling* scheme. The color coding is green for the theoretical diffusion PDF, maroon for the isotropic, red for the half-Gaussian, yellow for the bending fibre, black for the branching and blue for the the usual DTI prolate model. [Maybe better to have a legend for the colors? Thicker lines?]

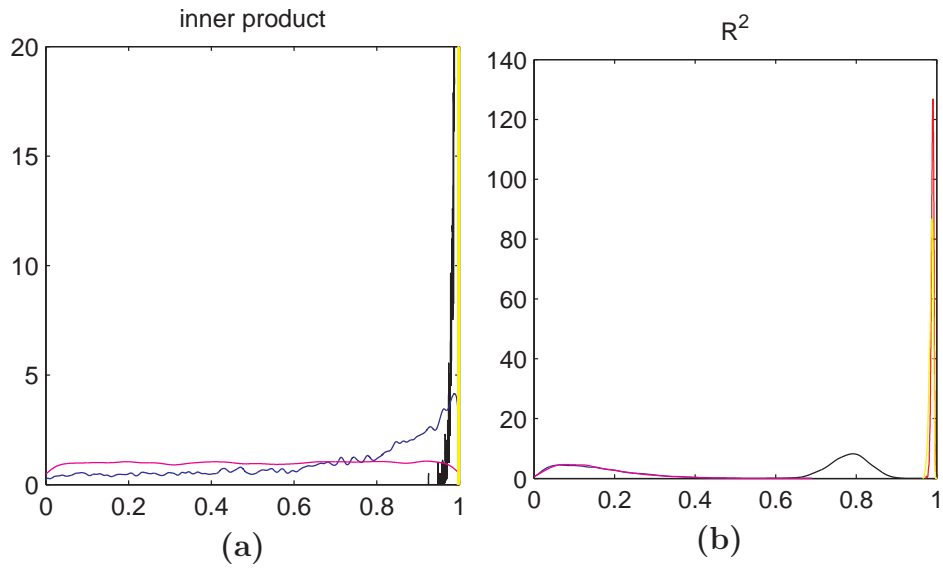


Figure 7: Estimating the linear phase term under the sparse sampling scheme: **(a)** the norm of $\mathbf{d}^{(\text{im})}$ and **(b)** the coefficient of determination R^2 from the linear regression. The two real-valued diffusions (prolate, blue) and (isotropic, purple), forking fibre (black), a single asymmetric diffusion (red), or bending fibre (yellow).

and 3(b,i), it is apparent that the main structure of Figs. 1(a,i) and 1(b,i) may be described as linear, but Figs. 3(a,i) and 3(b,i) are clearly not linear. Fig. 3(a,i) exhibits additional asymmetry, while Fig. 3(b,i) has the additional double maxima. Such empirical observations are born out by the distributional results in Fig. 7(b). **[More discussion about Figure 6!!!]**

4.2. Clinical data

We show the main directions of the real magnitude (associated with the largest eigenvalue) and the gradient of the estimated linear phase in Figs. 8(a)–(d), 9(a)–(d), and 10(a)–(d), all weighted by the FA. In all three figures we have used either (c) a p -value of 0.1 to determine whether average phase is zero, combined with the FDR procedure with $p = 0.01$ on the gradient of the phase in the retained voxels, or (d) no check on the average phase and a non-adjusted p -value of 0.1 for testing the gradient of the phase. We show the magnitude of the phase vector (Figs. 11(a,d,g)) and the R^2 associated with fitting the linear phase factor (Figs. 11(b,e,h)), as well as the estimated model order (Figs. 11(c,f,i)). We apply the false discovery rate (FDR) procedure for multiple comparisons (Benjamini and Hochberg, 1995; Genovese et al., 2002) to determine which voxels have an average phase equal to zero, but exhibit linear structure (Fig. 12). It would appear that we have identified some bending structure, bending left/right. Note that the three-dimensional plots have been rotated for better view (consistently per voxel), and so the phase structure in Fig. 13 does not admit direct interpretation from the image (see instead 12(f)). **[Contrast Figure 12a with Fig. 13 in Jones (2008).]**

5. Discussion

What have we done? There is phase information in DW-MRI. This information is “free”, and neglecting the information leads to inefficient estimation.

When the spurious phase has been removed, structure remains. This has to do with overall diffusion asymmetries, and allows us to recognize bending and (hypothetically) forking structure. In the case of bending, it is anticipated that a linear approximation to the phase would be sufficient, and the data does confirm this to be the case, with an observed R^2 of nearly 0.8 in some cases.

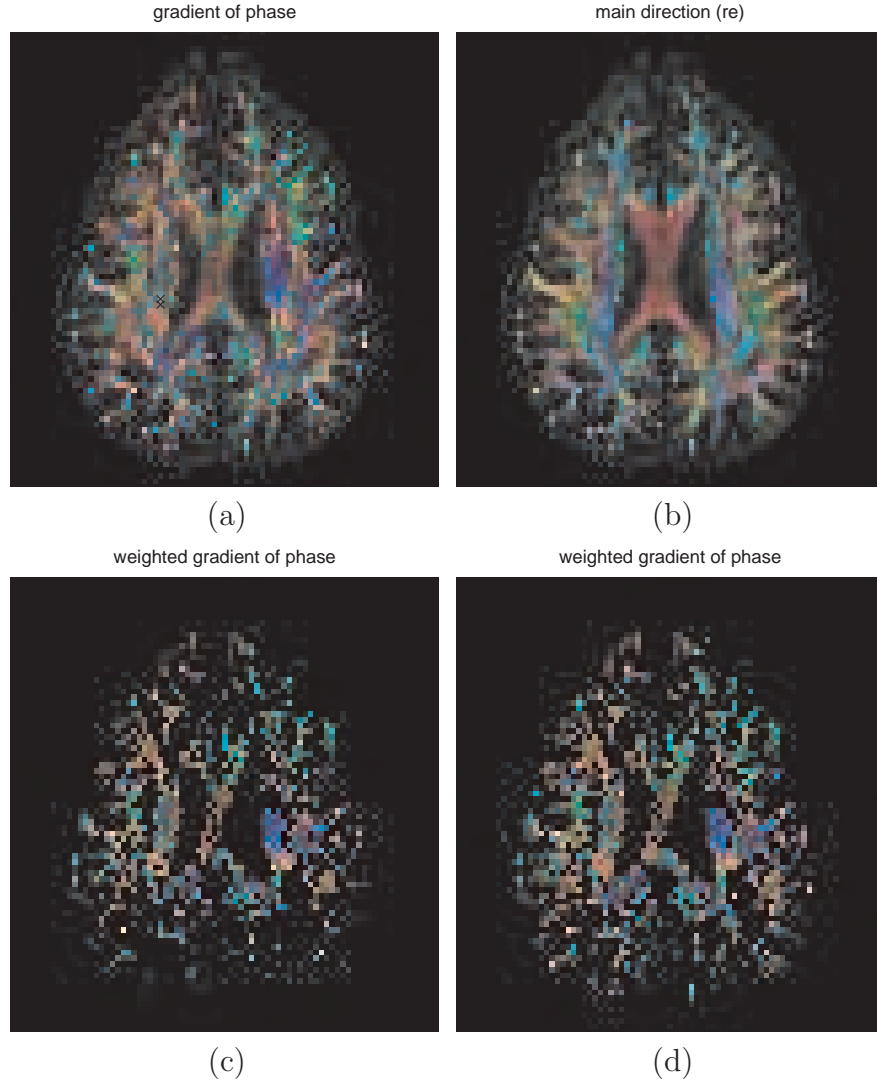


Figure 8: Estimated orientations from *in vivo* data using the (a) gradient of the phase and (b) the dominant direction of the amplitude, (c) and (d) showing only directions that may be considered statistically significant (see the text for details). This is slice 10.

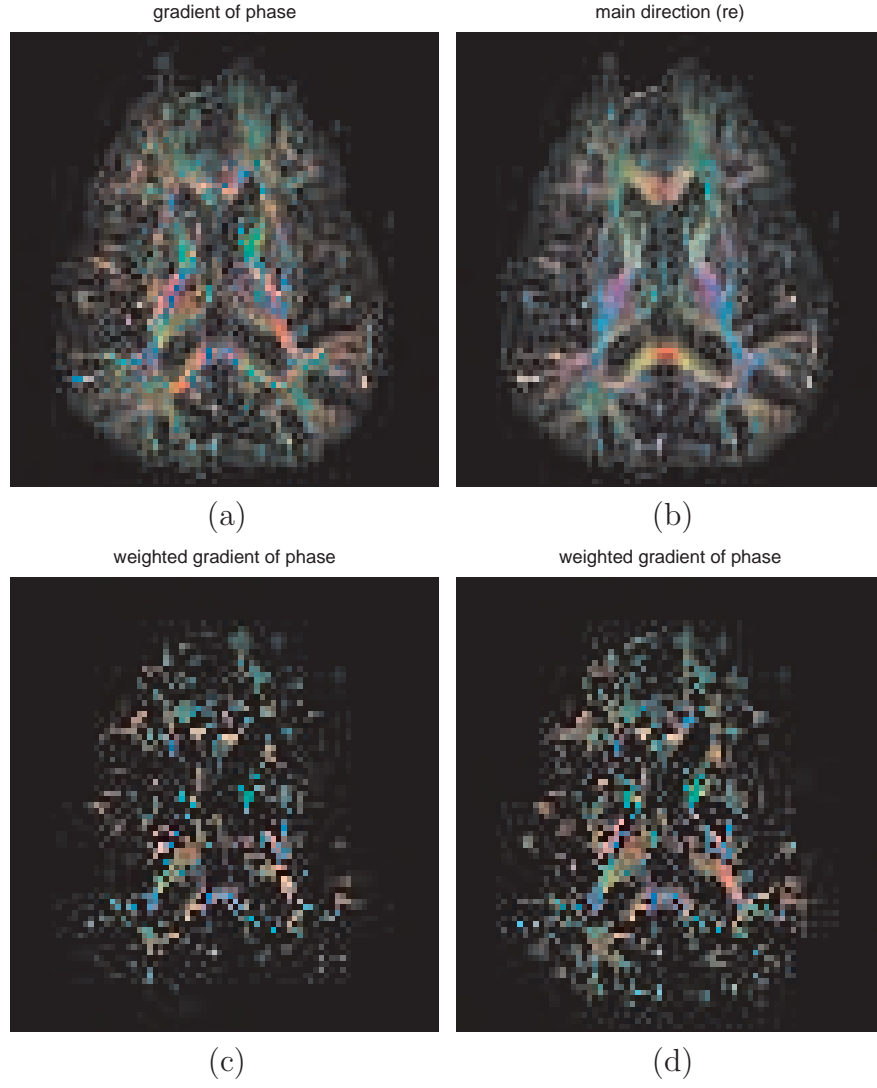


Figure 9: Estimated orientations from *in vivo* data using the (a) gradient of the phase and (b) the dominant direction of the amplitude, (c) and (d) showing only directions that may be considered statistically significant (see the text for details). This is slice 7.

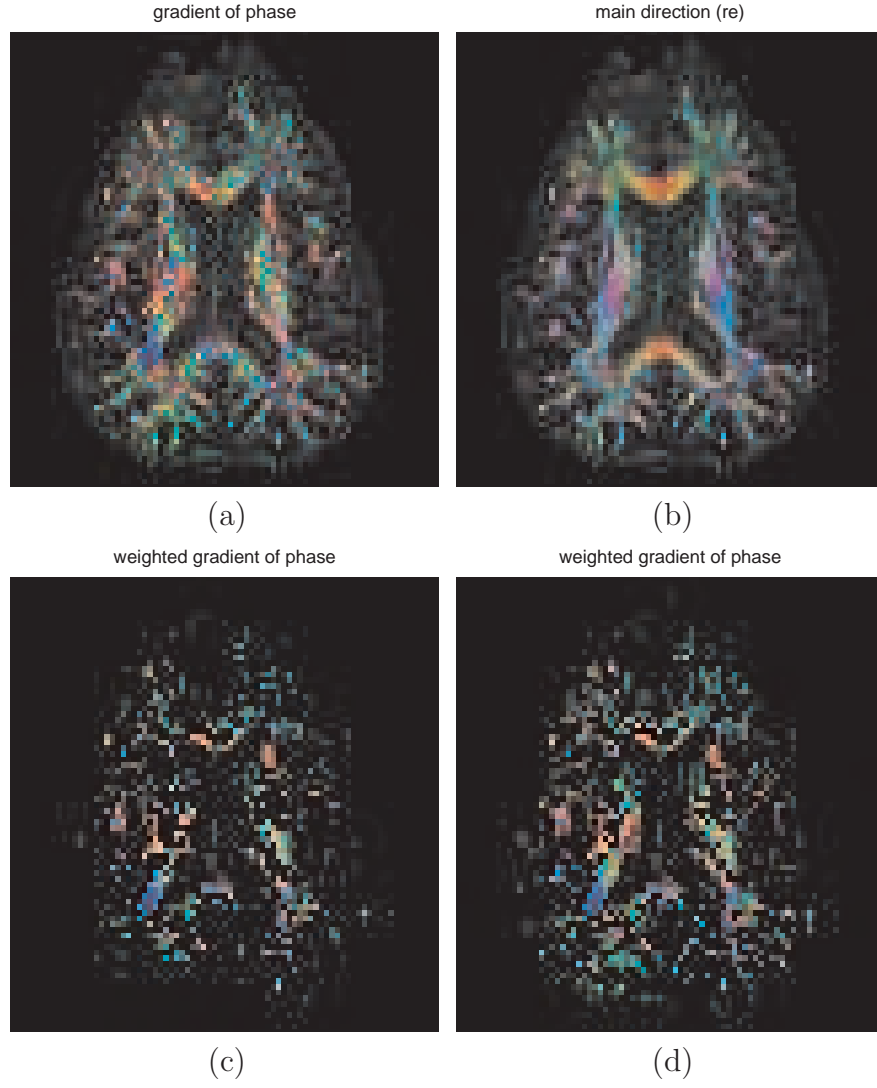


Figure 10: Estimated orientations from *in vivo* data using the (a) gradient of the phase and (b) the dominant direction of the amplitude, (c) and (d) showing only voxels where these may be considered statistically significant (see the text for details).

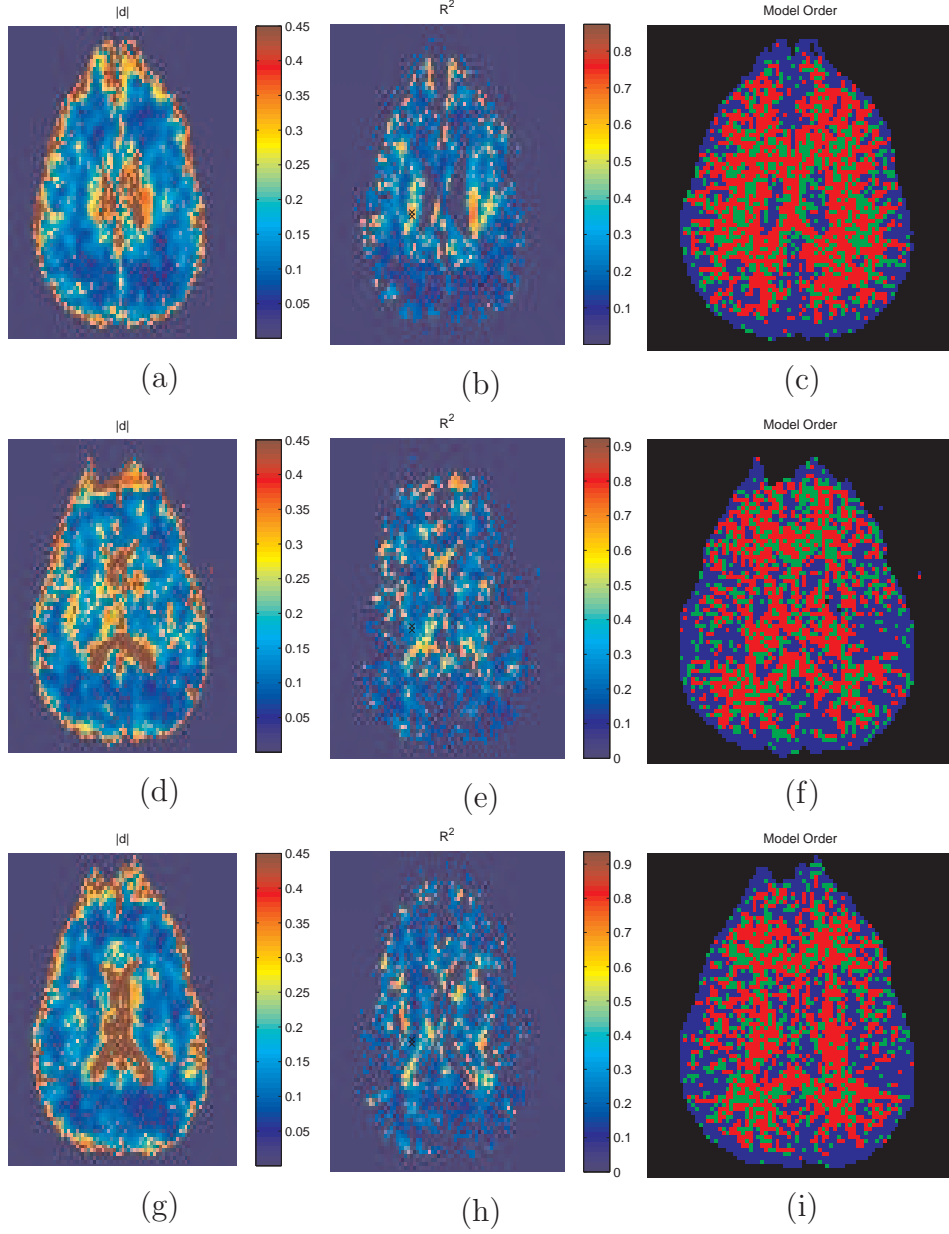


Figure 11: (a, d and g) the estimated norm for the phase vector, (b, e and g) the goodness-of-fit value R^2 weighted by a one-zero mapping based on the FA for interpretability and (c, f, i) the estimated model order. The first row is slice 10, the second row is slice 7 and the third row is slice 8.

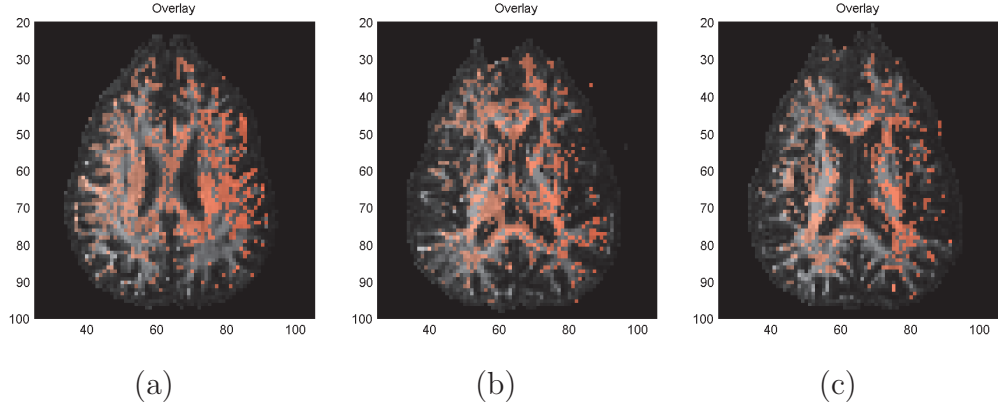


Figure 12: Three slices from the clinical data acquisition, where voxels highlighted in red provide indicate strong Hermitian structure and the average phase is zero.

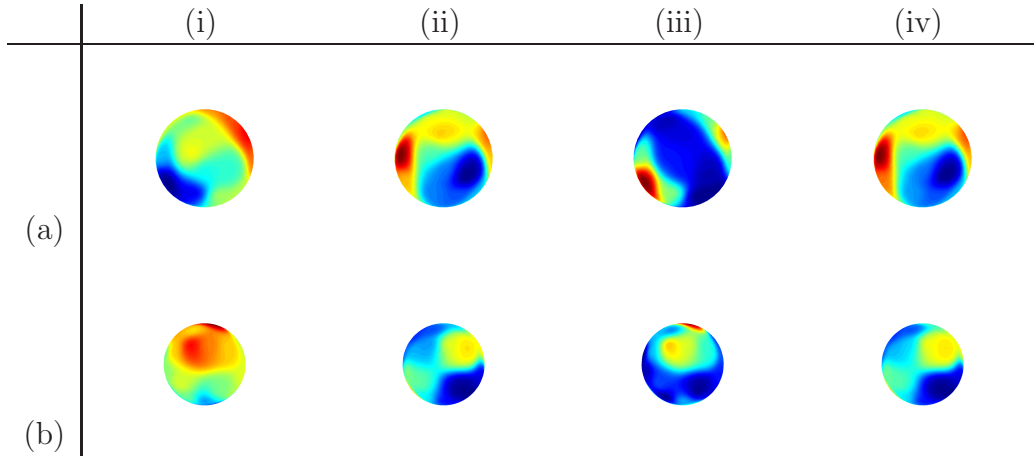


Figure 13: Two voxels of clinical data from the same axial slice in Fig. 12. For both voxels the columns denote the (i) phase, (ii) magnitude of the real part of the complex-valued signal, (iii) the magnitude of the imaginary part of the complex-valued signal and (iv) the amplitude of the complex-valued signal. For both voxels provided here the phase displays asymmetry similar to ???, with an $R^2 = 0.64$.

What does it do? This information identifies local symmetries such as bending or forking, and we can locally describe bending well.

What is usually done? The Gaussian diffusion model has greatly facilitated automated analysis of DTI data. Naturally, there are limitations in the types of white-matter microstructure that can be described by the Gaussian diffusion model. Crossing fibers, and other higher-order structures, do not agree with the simple model and produce mixed results in terms of common scalar statistical summaries; e.g., fractional anisotropy. Methodological developments provide a variety of methods and models beyond the Gaussian diffusion model; e.g., tensor mixtures, DOT, PAS-MRI, spherical deconvolution, etc. A line of investigation has looked at diffusion kurtosis, and higher order properties of the PDF; e.g., diffusion kurtosis imaging (Jensen et al., 2005).

Another possible non-Gaussian structure is asymmetry, associated with bending and forking. A Gaussian distribution is always symmetric, whilst some typical features of a PDF cannot be encompassed by a symmetric distribution. To detect asymmetries it is necessary to collect complex-valued observations.

5.1. The importance of pre-processing

The analysis of such observations is unfortunately not trivial: initially various spurious phase effects must be removed. Such removal can be unsuccessful in which case subsequent analysis becomes highly questionable. It is therefore necessary to start after such effects have been removed by testing if the remaining phase structure is Hermitian symmetric.

5.2. Strengths and weaknesses

However, single-shell data is unlikely to yield sufficient phase information, and as when employing higher-order tensor models, more data are needed to fit the models. **We have done the best with a single shell of data!**

5.3. Applications

The phase information can be used to improve parametric DW-MRI analysis methods.

With phase information, various tracking algorithms could be improved, in a similar manner to using the scalene structure of diffusions (Seunarine et al., 2007). The advantage of the complex-valued information is that we get an estimated orientation of the bending, and this can be feed directly into the

tracking algorithm. For forking structure a simple linear phase will not do, see Fig. 1. In this case we anticipate a small value of R^2 will indicate voxels in which the assumption of linear phase will be insufficient. We cannot however fit such models from a single shell, and must collect more complicated data to be able to deduce such structure.

Acknowledgements

The authors would like to thank...

Appendix A. Data pre-processing

Appendix A.1. Individual voxel/gradient/coil model

We write the observed signal at spatial location \mathbf{x}_n , coil γ and diffusion sensitisation gradient \mathbf{q}_j (unit vector) as

$$Z_\gamma(\mathbf{q}_l; \mathbf{x}_n) = X_\gamma(\mathbf{q}_l; \mathbf{x}_n) - iY_\gamma(\mathbf{q}_l; \mathbf{x}_n) \quad (\text{A.1})$$

$$= A_\gamma(\mathbf{q}_l; \mathbf{x}_n) s_\gamma(\mathbf{q}_l; \mathbf{x}_n) \exp(-i\phi_\gamma(\mathbf{q}_l; \mathbf{x}_n)) \quad (\text{A.2})$$

$$= \mathcal{Z}_\gamma(\mathbf{q}_l; \mathbf{x}_n) + \epsilon_\gamma(\mathbf{q}_l; \mathbf{x}_n) \quad (\text{A.3})$$

where $\gamma = 1, \dots, n_\gamma$ and $l = 1, \dots, L$. We refer to $A_\gamma(\mathbf{q}_l; \mathbf{x}_n)$ as the observed amplitude, $s_\gamma(\mathbf{q}_l; \mathbf{x}_n)$ is the observed coil sensitivity and $\phi_\gamma(\mathbf{q}_l; \mathbf{x}_n)$ is the observed phase. These are also unfortunately noisy observations, contaminating the noise-free observations

$$\mathcal{Z}_\gamma(\mathbf{q}_l; \mathbf{x}_n) = \mathcal{A}(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma(\mathbf{q}_l; \mathbf{x}_n) \exp(-i\varphi_\gamma(\mathbf{q}_l; \mathbf{x}_n)) \quad (\text{A.4})$$

by Gaussian proper noise (Picinbono, 1996), $\epsilon_\gamma(\mathbf{q}_l; \mathbf{x}_n)$. We refer to $\mathcal{A}(\mathbf{q}_l; \mathbf{x}_n)$ as the (true) amplitude, $\varsigma_\gamma(\mathbf{q}_l; \mathbf{x}_n)$ as the (true) coil sensitivity and $\varphi_\gamma(\mathbf{q}_l; \mathbf{x}_n)$ as the (true) phase. The IFT of $\mathcal{Z}_\gamma(\mathbf{q}_l; \mathbf{x}_n)$ is denoted $a_\gamma(\mathbf{x}; \mathbf{x}_n)$. We model the variance as given by

$$\text{Var}\{\epsilon_\gamma(\mathbf{q}_l; \mathbf{x}_n)\} = \sigma^2(\mathbf{x}_n), \quad \forall \gamma, l. \quad (\text{A.5})$$

This means that we model the variance as evolving spatially but as staying fixed across coils γ and orientations \mathbf{q}_l . This is equivalent to considering the main source of noise as physiological rather than thermal (thermal noise is known to be coil dependent). We also assume that $\{\gamma\}$ has been chosen so that

$$\sum_{\gamma} \varsigma_\gamma(\mathbf{q}_l; \mathbf{x}_n) = C, \quad (\text{A.6})$$

some constant.

Appendix A.2. Modelling rigid motion

With the model of Section Appendix A.1, we intend to determine the motion artifacts and remove these from the observations. We start by processing the observations at a fixed slice position x_3 and fixed coil γ as well as gradient direction l , thus only dealing with functions in the plane of (x_1, x_2) . If we shift the image by $\mathbf{x}^{(o)} = (x_1^{(o)}, x_2^{(o)})^T$ in the (x_1, x_2) plane then the \mathbf{q} space representation changes by

$$\begin{aligned} \mathcal{Z}_\gamma^{\mathbf{x}^{(o)}}(\mathbf{q}_l; \mathbf{x}_n) &= \int \int \int a_\gamma(\mathbf{y} - \mathbf{x}^{(o)}; \mathbf{x}_n) \exp(-2i\pi \mathbf{q}_l^T \mathbf{y}) d^3\mathbf{y} \\ &= \exp(-2i\pi \mathbf{q}_l^T \mathbf{x}^{(o)}) \mathcal{Z}_\gamma(\mathbf{q}_l; \mathbf{x}_n). \end{aligned} \quad (\text{A.7})$$

$\mathbf{x}^{(o)}$ is implicitly a function of \mathbf{x}_n . Thus the motion will imply a linear phase shift by $\mathbf{x}^{(o)}$. However the motion is not quite rigid, and additionally motion will happen in the full x_1, x_2 and x_3 space, not just the (x_1, x_2) plane. We therefore deduce that the motion will approximately have the effect of the non-linear shift operator $\mathcal{S}\{\}$ (Olhede, 2008)

$$\mathcal{S}\{\mathcal{Z}_\gamma\}(\mathbf{q}_l; \mathbf{x}_n) = \exp(-2i\pi \overline{\varphi}_{\gamma l}(\mathbf{x}_n)) \mathcal{Z}_\gamma(\mathbf{q}_l; \mathbf{x}_n), \quad (\text{A.8})$$

where $\overline{\varphi}_{\gamma l}(\mathbf{x}_n)$ is some smooth function in \mathbf{x}_n . The observed motion induced phase is in fact *mainly* smooth with some jumps due to the wrap-round of the 2π modulus observation of the $\tan^{-1}(\cdot)$ function. Note that we model large scale spatial effects using \mathbf{x}_n , while the small-scale spatial variable is \mathbf{x} the canonical variable of \mathbf{q} .

Appendix A.3. Estimating rigid motion phase

To estimate the motion induced phase we therefore look at Eq. (A.8) which suggest that the low frequency structure of the observed data is subject to the same phase motion as the high-frequency structure. It therefore is reasonable to determine this phase from the low-frequency structure of the data, and we define with some specified smoothness level α a projection of the signal to low frequencies

$$\begin{aligned} \mathcal{P}_\alpha\{\mathcal{Z}_\gamma\}(\mathbf{q}_l; \mathbf{x}_n) &= \overline{A}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \overline{s}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \exp\left(-i\overline{\phi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n)\right) \\ &= \overline{Z}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n). \end{aligned}$$

This projection is associated with a fixed bandwidth α , that we shall assume *known*, but is de-facto user tuned in this paper. We define an estimator of the true phase as

$$\hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) = \phi_\gamma(\mathbf{q}_l; \mathbf{x}_n) - \overline{\phi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n). \quad (\text{A.9})$$

That is, at each coil, direction and location \mathbf{x}_n we define a high frequency phase as being the observed phase minus the phase of the low-frequency filtered observation $\overline{Z}_\gamma(\mathbf{q}_l; \mathbf{x}_n)$. The advantage of this model is that we avoid a *parametric* model for $\overline{\phi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n)$.

Appendix A.4. Properties of rigid motion phase

We shall make the assumption that α has been chosen appropriately and so with $\overline{\varphi}_{\gamma l}(\mathbf{x}_n)$ denoting the smooth motion induced phase and $\varphi(\mathbf{q}_l; \mathbf{x}_n)$ the observation excess phase

$$\mathbb{E} \{ \phi_\gamma(\mathbf{q}_l; \mathbf{x}_n) \} = \overline{\varphi}_{\gamma l}(\mathbf{x}_n) + \varphi(\mathbf{q}_l; \mathbf{x}_n) \quad (\text{A.10})$$

$$\mathbb{E} \{ \overline{\phi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \} = \overline{\varphi}_{\gamma l}(\mathbf{x}_n) \quad (\text{A.11})$$

$$\mathbb{E} \{ \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \} = \varphi(\mathbf{q}_l; \mathbf{x}_n), \quad (\text{A.12})$$

if the smoothing parameter α has been chosen correctly, so that the procedure neither over-, nor under-smooths. We wish to estimate $\varphi(\mathbf{q}_l; \mathbf{x}_n)$, and thus need to combine the γ coil estimates over $\gamma = 1, \dots, n_\gamma$. This will depend on the properties of $\hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n)$ and we determine that

$$\begin{aligned} \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) &= -\Im \{ \ln (Z_\gamma(\mathbf{q}_l; \mathbf{x}_n)) \} + \Im \{ \ln (\overline{Z}_\gamma(\mathbf{q}_l; \mathbf{x}_n)) \} \\ &= -\Im \left\{ \ln \left(\mathcal{A}(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma(\mathbf{q}_l; \mathbf{x}_n) \exp \left(-i\overline{\varphi}_{\gamma l}(\mathbf{x}_n) - i\overline{\varphi}(\mathbf{q}_l; \mathbf{x}_n) \right) + \epsilon_\gamma(\mathbf{q}_l; \mathbf{x}_n) \right) \right\} \\ &\quad + \Im \left\{ \ln \left(\overline{\mathcal{A}}(\mathbf{q}_l; \mathbf{x}_n) \overline{\varsigma}_\gamma(\mathbf{q}_l; \mathbf{x}_n) \exp \left(-i\overline{\varphi}_{\gamma l}(\mathbf{x}_n) \right) + \overline{\epsilon}_\gamma(\mathbf{q}_l; \mathbf{x}_n) \right) \right\}. \end{aligned} \quad (\text{A.13})$$

It then follows that (assuming the SNR is sufficiently large), with HOT denoting higher order terms,

$$\begin{aligned}
\hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) &= \bar{\varphi}_{\gamma l}(\mathbf{x}_n) + \varphi(\mathbf{q}_l; \mathbf{x}_n) - \bar{\varphi}_{\gamma l}(\mathbf{x}_n) \\
&\quad - \Im \left\{ \frac{\epsilon_\gamma(\mathbf{q}_l; \mathbf{x}_n)}{\exp \left(-i(\bar{\varphi}_{\gamma l}^{(\alpha)}(\mathbf{x}_n) + \varphi(\mathbf{q}_l; \mathbf{x}_n)) \right) \mathcal{A}(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma(\mathbf{q}_l; \mathbf{x}_n)} \right. \\
&\quad \left. - \frac{\bar{\epsilon}_\gamma(\mathbf{q}_l; \mathbf{x}_n)}{\exp \left(-i\bar{\varphi}_{\gamma l}^{(\alpha)}(\mathbf{x}_n) \right) \bar{\mathcal{A}}(\mathbf{q}_l; \mathbf{x}_n) \bar{\varsigma}_\gamma(\mathbf{q}_l; \mathbf{x}_n)} \right\} + \text{HOT} \\
&= \varphi(\mathbf{q}_l; \mathbf{x}_n) + \tilde{\epsilon}(\mathbf{q}_l; \mathbf{x}_n) + \text{HOT}, \tag{A.14}
\end{aligned}$$

and the higher-order terms may be ignored (with sufficiently large SNR).

Appendix A.5. Properties of the estimated phase

Thus it follows that with L^2 being the area covered by the low pass filtered version of the signal

$$\mathbb{E} \{ \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \} = \varphi(\mathbf{q}_l; \mathbf{x}_n) + \text{HOT}, \tag{A.15}$$

and so

$$\begin{aligned}
\text{Var} \{ \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \} &= \frac{\sigma^2(\mathbf{x}_n)}{\mathcal{A}^2(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma^2(\mathbf{q}_l; \mathbf{x}_n)} + \frac{\sigma^2(\mathbf{x}_n)}{L^2 \mathcal{A}^2(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma^2(\mathbf{q}_l; \mathbf{x}_n)} \\
&\quad - 2 \Re \left\{ \frac{\sigma^2(\mathbf{x}_n)}{L \exp(i\varphi(\mathbf{q}_l; \mathbf{x}_n)) \mathcal{A}^2(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma^2(\mathbf{q}_l; \mathbf{x}_n)} \right\} \\
&= \frac{\sigma^2(\mathbf{x}_n)}{\mathcal{A}^2(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma^2(\mathbf{q}_l; \mathbf{x}_n)} \left(1 + \frac{1}{L^2} - 2 \frac{\cos(\varphi(\mathbf{q}_l; \mathbf{x}_n))}{L} \right).
\end{aligned}$$

We can deduce directly from the equation above that

$$\begin{aligned}
\frac{\sigma^2(\mathbf{x}_n)}{\mathcal{A}^2(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma^2(\mathbf{q}_l; \mathbf{x}_n)} \left[1 - \frac{1}{L} \right]^2 &\leq \text{Var} \{ \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \} \\
&\leq \frac{\sigma^2(\mathbf{x}_n)}{\mathcal{A}^2(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma^2(\mathbf{q}_l; \mathbf{x}_n)} \left[1 + \frac{1}{L} \right]^2.
\end{aligned}$$

While we could keep the exact form – such as given above – for simplicity we shall approximate the variance to be:

$$\text{Var} \{ \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \} \sim \frac{\sigma^2(\mathbf{x}_n)}{\mathcal{A}^2(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma^2(\mathbf{q}_l; \mathbf{x}_n)}. \tag{A.16}$$

Appendix A.6. Estimating excess phase and its properties

To avoid a deluge of symbols in the subsequent calculations we take

$$\mathcal{A}_\gamma(\mathbf{q}_l; \mathbf{x}_n) = \mathcal{A}(\mathbf{q}_l; \mathbf{x}_n) \varsigma_\gamma(\mathbf{q}_l; \mathbf{x}_n), \quad (\text{A.17})$$

this giving us one amplitude at each coil and orientation. We also have assumed that

$$\text{Cov} \left\{ \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n), \hat{\varphi}_{\gamma'}^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n) \right\} = 0, \quad \text{if } \gamma \neq \gamma', \quad (\text{A.18})$$

or that information from different coils is independent. We then define aggregate (over γ) estimators at each gradient orientation by

$$\hat{\mathcal{A}}_\gamma^{(n)}(\mathbf{q}_l; \mathbf{x}_n) = |Z_\gamma(\mathbf{q}_l; \mathbf{x}_n)|, \quad (\text{A.19})$$

$$A(\mathbf{q}_l; \mathbf{x}_n) = \sum_\gamma \hat{\mathcal{A}}_\gamma^{(n)}(\mathbf{q}_l; \mathbf{x}_n) \quad (\text{A.20})$$

$$\phi(\mathbf{q}_l; \mathbf{x}_n) = \frac{\sum_\gamma \hat{\mathcal{A}}_\gamma^{(n)}(\mathbf{q}_l; \mathbf{x}_n) \hat{\varphi}_\gamma^{(\alpha)}(\mathbf{q}_l; \mathbf{x}_n)}{\sum_\gamma \hat{\mathcal{A}}_\gamma^{(n)}(\mathbf{q}_l; \mathbf{x}_n)} \quad (\text{A.21})$$

$$\text{E} \{ \phi(\mathbf{q}_l; \mathbf{x}_n) \} = \varphi(\mathbf{q}_l; \mathbf{x}_n) + \text{HOT}$$

$$\text{Var} \{ \phi(\mathbf{q}_l; \mathbf{x}_n) \} = \frac{\sigma^2(\mathbf{x}_n)}{\left[\sum_\gamma \mathcal{A}_\gamma^{(n)}(\mathbf{q}_l; \mathbf{x}_n) \right]^2} + \text{HOT}.$$

The higher-order terms become important when the SNR is low. As long as the signal to noise ratio is sufficiently high, we find that $\phi(\mathbf{q}_l; \mathbf{x}_n)$ is approximately Gaussian. We can then define

$$Z(\mathbf{q}_l; \mathbf{x}_n) = A(\mathbf{q}_l; \mathbf{x}_n) \exp(-i\phi(\mathbf{q}_l; \mathbf{x}_n)), \quad (\text{A.22})$$

and this estimate will subsequently be treated like an observed $Z(\mathbf{q}_l; \mathbf{x}_n)$ (Eq. (1)) and the latter argument suppressed for brevity and additional clarity.

Abramowitz, M., Stegun, I. A., 1972. Handbook of Mathematical Functions, 10th ed. National Bureau of Standards.

Aksoy, M., Liu, C., Moseley, M. E., Bammer, R., 2008. Single-step nonlinear diffusion tensor estimation in the presence of microscopic and macroscopic motion. *Magnetic Resonance in Medicine* 59, 1138–1150.

- Alexander, D. C., 2005. Multiple-fibre reconstruction algorithms for diffusion MRI. *Annals of the New York Academy of Sciences* 1046, 113–133.
- Basser, P. J., Mattiello, J., Le Bihan, D., 1994a. MRI diffusion tensor spectroscopy and imaging. *Biophysical Journal* 66, 259–267.
- Basser, P. J., Mattiello, J., Le Bihan, D., 1994b. Estimation of the effective self-diffusion tensor from the NMR spin-echo. *Journal of Magnetic Resonance B* 103, 247–254.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical approach to multiple testing. *JRSS b* 57, 289–300.
- Bodurka, J., Jesmanowicz, A., Hyde, J. S., Xu, H., Estkowski, L., Li, S. J., 1999. Current-induced magnetic resonance phase imaging. *Journal of Magnetic Resonance* 137, 265–271.
- Bretthorst, G. L., 2006. *Bayesian Analysis Users Guide*.
- Callaghan, P. T., 1993. *Principles of Nuclear Magnetic Resonance Microscopy*. Clarendon Press, Oxford, UK.
- Chung, H.-W., Chou, M.-C., Chen, C.-Y., 2011. Principles and limitations of computational algorithms in clinical diffusion tensor MR tractography. *American Journal of Neuroradiology* 32, 3–13.
- Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R., 2007. Regularized, fast and robust analytical q -ball imaging. *Magnetic Resonance in Medicine* 58, 497–510.
- Genovese, C. R., Lazar, N. A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15, 870–878.
- Granlund, G. H., Knutsson, H., 1995. *Signal Processing for Computer Vision*. Kluwer, Dordrecht, The Netherlands.
- Hess, C. P., Mukherjee, P., Han, E. T., Xu, D., Vigneron, D. B., 2006. Q-ball reconstruction of multimodal fibre orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine* 56, 104–117.

- Hoaglin, D. C., Mosteller, F., Tukey, J. W., 1983. Understanding Robust and Exploratory Data Analysis. John Wiley & Sons, New York, USA.
- Hui, E. S., Cheung, M. M., Qi, L., Wu, E. X., 2008. Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis. *NeuroImage* 42, 122–134.
- Jansons, K. M., Alexander, D. C., 2003. Persistent angular structure: New insights from diffusion magnetic resonance imaging data. *Inverse Problems* 19, 1031–1046.
- Jenkinson, M., Smith, S. M., 2001. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* 5 (2), 143–156.
- Jensen, J. H., Helpert, J. A., Ramani, A., Lu, H., Kaczynski, K., 2005. Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine* 53, 1432–1440.
- Jones, D. K., 2008. Studying connections in the living human brain with diffusion MRI. *Cortex* 44, 936–952.
- Jones, D. K., Horsfield, M. A., Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine* 42, 515–525.
- Lazar, M., 2010. Mapping brain anatomical connectivity using white matter tractography. *NMR in Biomedicine* 23, 821–835.
- Lazar, M., Jensen, J. H., Xuan, L., Helpert, J. A., 2008. Estimation of the orientation distribution function from diffusional kurtosis imaging. *Magnetic Resonance in Medicine* 60, 774–781.
- Liu, C., Bammer, R., Acar, B., Moseley, M. E., 2004. Characterizing non-Gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine* 51, 924–937.
- Liu, C., Bammer, R., Moseley, M. E., 2005a. Limitations of apparent diffusion coefficient-based models in characterising non-Gaussian diffusion. *Magnetic Resonance in Medicine* 54, 419–428.

- Liu, C., Moseley, M. E., Bammer, R., 2005b. Simultaneous phase correction and SENSE reconstruction for navigated multi-shot DWI with non-cartesian k -space sampling. *Magnetic Resonance in Medicine* 54, 1412–1422.
- Lu, H., Jensen, J. H., Ramani, A., Helpert, J. A., 2006. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine* 19, 236–247.
- Menon, R. S., 2002. Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI. *Magnetic Resonance in Medicine* 47, 1–9.
- Newbould, R. D., Skare, S., Bammer, R., 2008. On the utility of complex-averages diffusion weighted images. In: *Sixteenth Annual Meeting of the International Society for Magnetic Resonance in Medicine*. ISMRM, Toronto, Canada, p. 1810.
- Olhede, S. C., 2008. Localisation of geometric anisotropy. *IEEE Transactions on Signal Processing* 56, 2133–2138.
- Özarslan, E., Mareci, T. M., 2003. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine* 50, 955–965.
- Picinbono, B., 1996. Second-order complex random vectors and normal distributions. *IEEE Transactions on Signal Processing* 44, 2637–2640.
- Rowe, D. B., 2005. Modeling both the magnitude and phase of complex-valued fMRI data. *NeuroImage* 25, 1310–1324.
- Rowe, D. B., Logan, B. R., 2004. Complex fMRI analysis with unrestricted phase is equivalent to a magnitude-only model. *NeuroImage* 24, 603–606.
- Rowe, D. B., Meller, C. P., Hoffmann, R. G., 2007. Characterizing phase-only fMRI data with an angular regression model. *Journal of Neuroscience Methods* 161, 331–341.
- Schwarz, G. E., 1978. Estimating the dimension of a model. *Annals of Statistics* 6, 461–464.

- Seunarine, K., Cook, P. A., Hall, M. G., Embleton, K. V., Parker, G. J. M., Alexander, D. C., 2007. Exploiting peak anisotropy for tracking through complex structure. In: *Mathematical Methods in Biomedical Image Analysis*. IEEE Computer Society, Rio de Janeiro, p. 1.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., Luca, M. D., Drobnjak, I., Flitney, D. E., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., Stefano, N. D., Brady, J. M., Matthews, P. M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23 (Supplement 1), 208–219.
- Truong, T. K., Chen, B., Song, A. W., 2008. Integrated SENSE DTI with corrections of susceptibility – and eddy current-induced geometric distortions. *NeuroImage* 40, 53–58.
- Tuch, D. S., 2004. Q-ball imaging. *Magnetic Resonance in Medicine* 52, 1358–1372.
- Wedeen, V. J., Hagmann, P., Tseng, W.-Y., Reese, T. G., Weisskoff, R. M., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine* 54, 1377–1386.
- Whitcher, B., Matthews, P. M., 2006. Noninvasive brain imaging for experimental medicine in drug discovery and development: promise and pitfalls. *International Journal of Pharmaceutical Medicine* 20 (3), 167–176.