Structured sparsity for spatially coherent fibre orientation estimation in diffusion MRI

A. Auría[a], A. Daducci[a,b], J.-P. Thiran[a,b], Y. Wiaux[c]

[a] *Signal Processing Lab (LTS5), École Polytechnique Fédérale de Lausanne, Switzerland*
[b] *University Hospital Center (CHUV) and University of Lausanne (UNIL), Switzerland*
[c] *Institute of Sensors, Signals & Systems, Heriot-Watt University, Edinburgh, UK*

Abstract

We propose a novel formulation to solve the problem of intra-voxel reconstruction of the fibre orientation distribution function (FOD) in each voxel of the white matter of the brain from diffusion MRI data. The majority of the state-of-the-art methods in the field perform the reconstruction on a voxel-by-voxel level, promoting sparsity of the orientation distribution. Recent methods have proposed a global denoising of the diffusion data using spatial information prior to reconstruction, while others promote spatial regularisation through an additional empirical prior on the diffusion image at each q-space point. Our approach reconciles voxelwise sparsity and spatial regularisation and defines a spatially structured FOD sparsity prior, where the structure originates from the spatial coherence of the fibre orientation between neighbour voxels. The method is shown, through both simulated and real data, to enable accurate FOD reconstruction from a much lower number of q-space samples than the state of the art, typically 15 samples, even for quite adverse noise conditions.

1. Introduction

The challenge in diffusion MRI is to infer features of the local tissue anatomy, composition and microstructure from water displacement measurements. Water diffusion in living tissues is highly affected by its cellular organization (Beaulieu, 2002). In particular, water does not diffuse equally in all directions in a highly ordered organ such as

Email address: `anna.auria@epfl.ch` (A. Auría)

the brain and this property can be exploited to study the structural neural connectivity in a non-invasive way. The estimation of fibre connectivity patterns *in vivo* represents a major goal in neuroscience but also in a clinical perspective, with applications for diagnosis of stroke, schizophrenia or Parkinson's disease. In order to reconstruct entire fibre pathways and hence brain connections, tractography algorithms nowadays rely on the orientations of maximal water diffusion in each voxel. Thus, an accurate reconstruction of the local fibre populations is crucial to ensure good performance of fibre-tracking.

A great variety of approaches have been proposed to tackle the problem of intra-voxel fibre orientation estimation. Diffusion Tensor Imaging (DTI) (Basser et al., 1994) is one of the simplest and fastest reconstruction techniques since it only requires sampling 6 points of the q-space. However, it is by construction unable to model multiple fibre populations within a voxel and thus it is not valid in regions with crossings. Diffusion Spectrum Imaging (DSI) (Wedeen et al., 2005), on the other hand, is a model-free imaging technique known to provide good imaging quality. Yet, it requires strong magnetic field gradients and long acquisition times, needing typically 256 samples for a good reconstruction. As a consequence, it generally becomes too time-consuming to be of real interest in a clinical perspective. Accelerated acquisitions, relying on as few sampling points as possible while still sensitive to fibre crossings represent thus a major goal in the field.

In the last years, *spherical deconvolution* (SD) methods (Tournier et al., 2004; Alexander, 2005; Tournier et al., 2007) have become very popular in the framework of local reconstruction since they can recover the fibre configuration with a relatively small number of points, typically from 30 up to 60. They consider that both anisotropy and magnitude of water diffusion in white matter (WM) are constant in the whole volume. Under this assumption, SD methods acknowledge the fact that the diffusion signal can be expressed as the convolution of a response function, or kernel, with the fibre orientation distribution function (FOD). The FOD is a real-valued function on the unit sphere that indicates the orientation and the volume fraction of the fibre populations in a voxel. The Constrained Spherical Deconvolution approach of Tournier et al. (2004, 2007) represents the first attempt to solve the ill-posed SD problem. It applies

Tikhonov regularisation, introducing a constraint on the ℓ_2 norm of the FOD, specially to ensure its positivity. Apart from the aforementioned work, most of the state-of-the-art methods to solve SD problems promote sparse regularisation based on ℓ_1 minimisation (Jian and Vermuri, 2007; Ramirez-Manzanares et al., 2007; Mani et al., 2014), where the ℓ_1 norm is defined, for any real vector, as the sum of the absolute value of its coefficients. Yet, Daducci et al. (2014b) acknowledge in recent work that ℓ_1 minimisation is formally inconsistent with the fact that the volume fraction sum up to unity, and demonstrate the superiority of ℓ_0-norm minimisation. All these local reconstruction methods solve the FOD recovery problem for each voxel independently and thus, do not exploit the spatial coherence of the fibre tracts in the brain. A number of approaches have addressed this shortcoming by formulating the problem globally (simultaneously for all voxels) to be able to exploit the correlation between the different volumes. Some of them decouple the problem and propose a global denoising of the diffusion data prior to reconstruction (Tristán-Vega and Aja-Fernández, 2010; Wiest-Daessl et al., 2008). Another group of methods present a joint scheme for reconstruction and spatial regularisation on the diffusion images at each q-space point. For instance, Fillard et al. (2007) propose a variational formulation to jointly estimate and regularise DTI to account for the effect of Rician noise in low SNR regimes, while Mani et al. (2014); Michailovich et al. (2011) use the standard state-of-the-art minimisation of the *total variation* (TV) semi-norm (Rudin et al., 1992) of the diffusion images.

In this paper, we propose a formulation that solves the fibre configuration of all voxels of interest simultaneously and imposes spatial regularisation directly on the fibre space. This reconstruction allows us to exploit information from the neighbouring voxels that cannot be taken into account by the existing state-of-the-art methods that approach fibre reconstruction independently in each voxel. The natural smoothness of the anatomical fibre tracts through the brain can be translated in a certain *spatial coherence* of the FOD in neighbouring voxels. Accordingly, in the aim of recovering the global FOD field in all voxels, the present work leverages a reweighted ℓ_1-minimisation scheme to promote a spatially structured sparsity prior imposing spatial coherence. While the spatial regularisation schemes proposed by Fillard et al. (2007); Mani et al. (2014); Michailovich et al. (2011) enforce sparsity of the images at each q-space point, our spatial regularisation relates to the fundamental coherence between fibre directions - the FOD - in neighbour voxels, thus adding anatomically driven constraints. Our code is available at https://github.com/basp-group/co-dmri and it is distributed open-source.

2. Materials and methods

2.1. dMRI framework for recovery of FOD via spherical deconvolution

In the SD framework, the intra-voxel structure estimation can be expressed through the FOD recovery problem in terms of the following linear formulation:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{\eta}, \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}_+^n$ stands for the FOD, $\boldsymbol{y} \in \mathbb{R}_+^m$ is the vector of measurements, Φ is the linear measurement operator and $\boldsymbol{\eta}$ is the acquisition noise. The reader can refer to Jian and Vermuri (2007) for a more detailed overview on SD methods and the formal equations describing the relationship between the FOD and the diffusion signal. We consider a dictionary Φ that spans a set of the Diffusion Basis Functions introduced by Ramirez-Manzanares et al. (2007). Each of these basis functions is generated by applying a different rotation to a kernel, which corresponds to the diffusion signal response to a single fibre. The set of available orientations represents a discretisation of half of the unit sphere ($\mathbb{S}^2$), assuming antipodal symmetry in diffusion signal. The diffusion signal can then be expressed as a linear combination of these basis functions, also referred to as the *atoms* of our dictionary Φ.

Prior constraints are essential to regularise a deconvolution problem like (1) in order to find a unique solution from an originally ill-posed problem. In the framework of the recently developed theory of *compressed sensing* (CS) (Donoho, 2006; Candès et al., 2006) sparsity priors are commonly used as regularisers to recover a signal from a set of undersampled measurements. In formulation (1) the sparsity can directly be inferred from the small number of fibre directions of interest, in correspondence with the FOD coefficients. In this paper, the method proposed by Daducci et al. (2014b) is taken as the state-of-the-art algorithm in the framework of SD local methods for FOD recovery. For the sake of completeness of this work, it is described in detail hereafter.

Daducci et al. (2014b) propose to resort explicitly to the non-convex ℓ_0 prior to solve for the FOD rather than to its convex ℓ_1 relaxation. A convex optimisation problem for FOD reconstruction can be defined through a constrained formulation between adequate sparsity prior and data, also making use of a reweighted sparse deconvolution. The proposed minimisation problem reads as:

$$\min_{\boldsymbol{x}\geq 0} \quad ||\Phi\boldsymbol{x}-\boldsymbol{y}||_2^2 \quad \text{s.t.} \quad ||\boldsymbol{x}||_0 \leq k. \tag{2}$$

In (2), $||\cdot||_0$ represents the ℓ_0 norm (number of non-zero coefficients) and k acts as a bound on the expected number of fibre populations in a voxel. Since the ℓ_0 norm is non-convex, a reweighted ℓ_1-minimisation scheme (Candès et al., 2008) is used in order to approach ℓ_0 minimisation by a sequence of convex weighted-ℓ_1 problems of the form:

$$\min_{\boldsymbol{x}\geq 0} \quad \|\Phi\boldsymbol{x}-\boldsymbol{y}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{x}\|_{w,1} \leq k. \tag{3}$$

In (3), the ℓ_0 norm has been substituted by a weighted-ℓ_1 norm defined as $\|\boldsymbol{x}||_{w,1} = \sum_i \boldsymbol{w}_i|\boldsymbol{x}_i|$. The algorithm alternates between estimating the solution at iteration t, $\boldsymbol{x}^{(t)}$, and redefining the weights essentially as the inverse of the values of the solution at the previous iteration $\boldsymbol{w}_i^{(t+1)} \approx 1/\boldsymbol{x}_i^{(t)}$. The use of these weights allows the algorithm to iteratively better estimate the non-zero locations and induces that, at convergence, the weighted-ℓ_1 norm mimics the ℓ_0 norm. Hence, formulation (2) promotes sparsity through a sequence of problems (3). In the rest of the manuscript we will refer to this voxel-by-voxel method based on ℓ_2 and ℓ_0 priors as `L2L0`.

In the next subsection we describe an algorithm, inspired by `L2L0`, that exploits the anatomical coherence of the fibre tracts of the brain by promoting a *structured sparsity* prior on the FOD field. We show evidence that taking into account neighbouring information through an appropriate prior directly on the object of interest improves significantly the results in comparison with solving for all voxels independently or using indirect spatial regularisation schemes.

2.2. Spatial regularisation through structured sparsity

In the aim of exploiting the spatial coherence of the fibres in the brain when recovering the local fibre configuration, we formulate a problem to solve the ensemble FOD field for all voxels simultaneously. To emphasize the fact that the minimisation problem (2) is formulated separately for each voxel of the brain, we can rewrite it using the following notation:

$$\min_{\boldsymbol{x}^{(v)} \geq 0} \quad ||\Phi \boldsymbol{x}^{(v)} - \boldsymbol{y}^{(v)}||_2^2 \quad \text{s.t.} \quad ||\boldsymbol{x}^{(v)}||_0 \leq k, \tag{4}$$

where $\boldsymbol{x}^{(v)} \in \mathbb{R}_+^n$ represents the real-valued FOD in the particular voxel indexed v. By concatenating all vectors $\boldsymbol{x}^{(v)}$ columnwise, one can build a matrix $\mathsf{X} \in \mathbb{R}_+^{n \times N}$, whose columns correspond to the FOD in each particular voxel. The elements of matrix X will be indexed as X_{dv}, each row d being associated with the atom of the dictionary oriented in direction indexed d, each column v being associated with voxel indexed v, $\mathsf{X}_{\cdot v} = \boldsymbol{x}^{(v)}$, as represented in Figure 3. N denotes the total number of voxels we want to recover the fibre configuration from. The rows of $\Phi\mathsf{X}$ represent the modelled diffusion images at each q-space point.

In our proposed formulation, a global data term is minimised adding a sparsity constraint that simultaneously promotes spatial coherence of the solution. Inspired by formulation (3), we adopt a procedure that consists in solving a sequence of problems of the form:

$$\min_{\mathsf{X} \in \mathbb{R}_+^{n \times N}} \quad \|\Phi \mathsf{X} - \mathsf{Y}\|_2^2 \quad \text{s.t.} \quad \|\mathsf{X}\|_{\mathsf{W},1} \leq K, \tag{5}$$

where the matrix $\mathsf{Y} \in \mathbb{R}^{m \times N}$ is formed by the concatenation of all N measurement column vectors: $\mathsf{Y}_{\cdot v} = \boldsymbol{y}^{(v)} \in \mathbb{R}^m$. The sensing matrix Φ is exactly the same as in (4) and $\| \cdot \|_{\mathsf{W},1}$ stands for a weighted ℓ_1 norm of a matrix defined as:

$$\|\mathsf{X}\|_{\mathsf{W},1} = \sum_{d,v} \mathsf{W}_{dv} |\mathsf{X}_{dv}|. \tag{6}$$

The following paragraphs are devoted to describe in detail the reweighting scheme and define the weighting matrix W.

In a reweighted-ℓ_1 scheme, large weights will progressively tend to *discourage* nonzero entries whereas small weights will *promote* nonzero entries in the solution. The weighting matrix W has the same dimension as X and each of its entries acts as a weight for the corresponding entry of X. The weights should still represent the inverse value of the associated entry at the previous iteration, so as to lead to an ℓ_0-norm prior at convergence. However, a strong spatial coherence prior can actually be promoted by adapting the computation of the weights as follows. Our definition of the weights is driven by the underlying anatomical assumption that fibre bundles in neighbouring voxels should have very close orientations as the trajectories are smooth (schematically represented in Figure 1). In terms of the FOD, this premise implies that neighbour voxels should bear *similar* directions.

To translate this idea into a mathematical formulation of the weights we start by formally defining the concept of *neighbourhood*. Since each atom of the dictionary represents a direction d on the half sphere, we define an *angular neighbourhood* $\mathcal{N}(d)$ for each of them composed by the *closest atoms* (in terms of angular distance). In our implementation we have considered a maximal angular distance of $15°$ to delimit the neighbourhood of each atom. Analogously, for each voxel v of the brain we define its *spatial neighbourhood* $\mathcal{N}(v)$ as the group of 26 voxels that share either a face, an edge or a vertex with the voxel of interest v, commonly referred to as the 26-adjacent neighbourhood (Huang et al., 1998). A visual representation of both $\mathcal{N}(d)$ and $\mathcal{N}(v)$ is shown in Figure 2. For convenience, we define $\overline{\mathcal{N}(d)} = d \cup \mathcal{N}(d)$ and $\overline{\mathcal{N}(v)} = v \cup \mathcal{N}(v)$, the neighbourhoods that include the central element. We then define the neighbourhood of an element X_{dv} as the entries of X at the intersection of rows d and all its neighbour directions, and columns v and all its neighbour voxels: $\overline{\mathcal{N}(dv)} = \{(d', v'); d' \in \overline{\mathcal{N}(d)}, v' \in \overline{\mathcal{N}(v)}\}$, as it is schematically represented in Figure 3.

At each iteration, every element of the weighting matrix W_{dv} is set as the inverse of an *average* of the absolute values that X takes in the neighbourhood of X_{dv} in the previous iteration:

$$\mathsf{W}_{dv}^{(t+1)} = \left[\tau^{(t)} + \frac{1}{|\overline{\mathcal{N}(v)}|} \sum_{d'v' \in \overline{\mathcal{N}(dv)}} |\mathsf{X}_{d'v'}^{(t)}| \right]^{-1}. \tag{7}$$

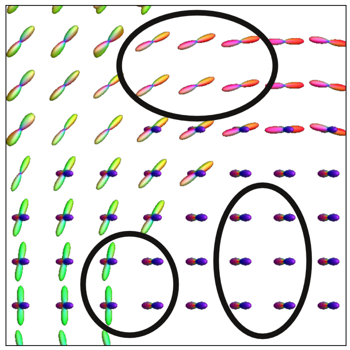

Figure 1: Synthetic FOD field in a representative 2D slice, which consists of two crossing fibre bundles. Due to the natural smoothness of the bundles, FODs in neighbouring voxels are expected to contain similar peaks, as highlighted in the figure.

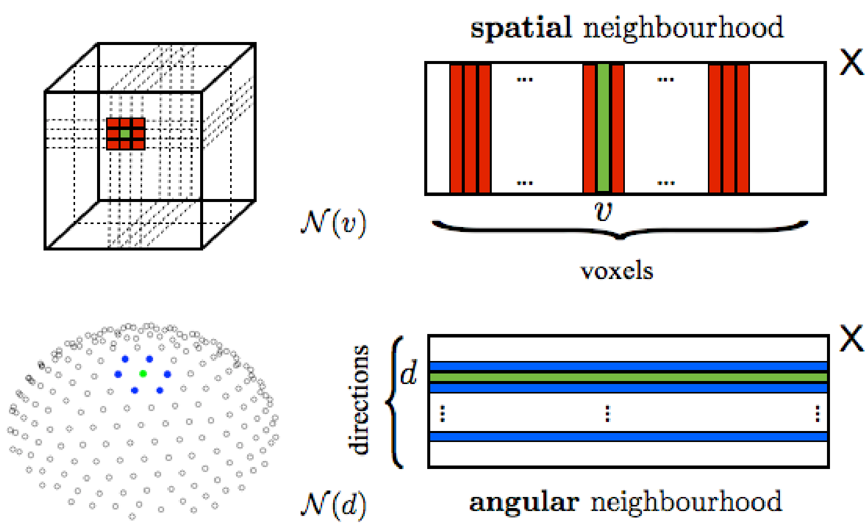

Figure 2: **Top row: Schematic representation of a spatial neighbourhood.** On the left: Set of voxels representing the 3D-volume (brain) we want to solve for. Voxels in red configure the neighbourhood $\mathcal{N}(v)$ for a particular voxel v, in green. On the right: Mapping of $\mathcal{N}(v)$ as a set of columns of matrix X. **Bottom row: Schematic representation of an angular neighbourhood.** On the left: Set of black circles representing the discretisation of the half sphere chosen to build dictionary Φ. Points highlighted in blue configure the neighbourhood $\mathcal{N}(d)$ for a particular direction d, in green. On the right: Mapping of $\mathcal{N}(d)$ as a set of rows of matrix X.

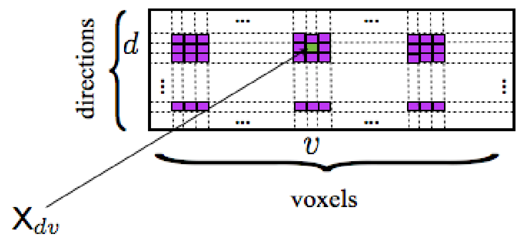

Figure 3: Schematic representation of the neighbourhood of element X_{dv} (in green), i.e. the elements of X involved in the computation of weight W_{dv}. It includes coefficients corresponding to directions d and all its neighbours $\mathcal{N}(d)$ for voxel v and all its neighbours $\overline{\mathcal{N}(dv)} = \{(d', v'); d' \in \overline{\mathcal{N}(d)}, v' \in \overline{\mathcal{N}(v)}\}$.

Consequently, at each iteration t, the weighting matrix $\mathsf{W}^{(t)}$ represents a *blurred* version of the current estimation of the solution $\mathsf{X}^{(t)}$[1]. In (7), we average over voxels, but sum over directions as all values in neighbour directions are interpreted as contributing to a single true local direction, in particular because the true direction does in general not coincide exactly to one of the discrete points of the sphere identifying our orientation dictionary. This helps to stabilise the regularisation and prevent the appearance of spurious peaks: fibre contributions are usually spread over a small angular support while spurious peaks are associated with isolated directions. To avoid infinite values for null averages, we add a stability parameter τ in the definition of the weights. We apply an homotopy strategy (Nocedal and Wright, 2006) and use a decreasing sequence $\{\tau^{(t)}\}$ in such a way that $\tau^{(t)} \to 0$ when $t \to \infty$. In the absence of any spatial constraint, $\mathsf{W}^{(0)}$ corresponds to the matrix of all 1s and thus, the weighted ℓ_1 norm is the standard $\ell_{1,1}$ norm of a matrix, $\|\mathsf{X}\|_{\mathsf{W},1} = \|\mathsf{X}\|_{1,1}$.

The specific computation of the weights described in the former paragraphs encourages that neighbour voxels present the same or very close (neighbour) directions, imposing *structured sparsity* of the solution. Indeed, all entries corresponding to the neighbourhood of an element contribute to its weight. Therefore those orientations

[1]The values of the final solution are influenced by their weights, however they are not directly identified with them.

that are “supported” by the surrounding voxels are reinforced, since they will be given a small weight compared to isolated directions that are not coherent with their environment. At convergence, our definitions (6) and (7) thus implement a spatially coherent version of the matrix ℓ_0 norm, i.e. the sum of the ℓ_0 norms of its columns. This reweighting scheme promotes a regularisation that takes into account the true anatomy of the brain accounting for the fact that fibre populations present a coherent trajectory across voxels close to each other in the brain volume. This prior constitutes a powerful constraint that cannot be exploited when solving the problem independently for each voxel, like in (4).

The main steps of the *reweighting scheme* are reported in algorithm 1; in the remaining of the manuscript we will refer to it as $\texttt{L2L0}_{\texttt{NW}}$, in reference to the described *neighbour weighted* scheme. The reweighting process stops when the relative variation between successive solutions $\|\mathsf{X}^{(t)} - \mathsf{X}^{(t-1)}\|_2 / \|\mathsf{X}^{(t-1)}\|_2$ is smaller than some bound or after the maximum number of iterations allowed is reached.

Algorithm 1 Reweighted ℓ_1 minimisation for global reconstruction of the FOD

Require: $\mathsf{Y} \in \mathbb{R}^{m\times N}$; $\Phi \in \mathbb{R}^{n\times m}$; K; ν; τ_{thr}; $N_{\max}$; $\mathcal{N}(d),\ d = 1,..,n$; $\mathcal{N}(v),\ v = 1,..,N$
Ensure: FOD $\mathsf{X} \in \mathbb{R}_+^{n\times N}$
  Initialise $t \leftarrow 0$; $\mathsf{X}^{(0)} = 0$; $\mathsf{W}^{(0)} \leftarrow 1$
  while $\rho > \nu$ and $t < N_{\max}$ **do**
    Solve:
    $\mathsf{X}^{(t)} \leftarrow \min_{\mathsf{X}\in\mathbb{R}_+^{n\times N}} \quad \|\Phi\mathsf{X} - \mathsf{Y}\|_2^2 \quad \text{s.t.} \quad \|\mathsf{X}\|_{\mathsf{W},1} \leq K$
    Update $\mathsf{W}^{(t+1)}$
    Update $\rho = \|\mathsf{X}^{(t)} - \mathsf{X}^{(t-1)}\|_2 / \|\mathsf{X}^{(t-1)}\|_2$
    $t \leftarrow t + 1$
  end while
  $\mathsf{X} \leftarrow \mathsf{X}^{(t-1)}$

2.3. Implementation details

To generate the dictionary Φ in our experiments, we estimated two different Gaussian kernels that model the diffusion signal in the regions of the brain corresponding to (i) white matter (WM) and (ii) partial volume with grey matter or cerebrospinal fluid (CSF). Modelling each kernel actually corresponds to estimating the three eigenvalues of the diffusion tensor. Grey matter and CSF are typically isotropic media. Consequently, their representative kernel is spherical – a tensor with three equal eigenvalues – and not sensitive to rotations. On the other hand, the kernel corresponding to the WM is anisotropic. Its response function was first estimated by fitting a tensor from the diffusion signal in those voxels with the highest fractional anisotropy (as expected to contain only one fibre population) and subsequently it was rotated in 200 different directions equally distributed on the sphere. Therefore, the final number of atoms of the dictionary used for this reconstruction is 201: 200 atoms corresponding to WM plus 1 isotropic atom modelling partial volume with CSF and grey matter.

Each weighted-ℓ_1 problem of the form (5) is solved using Douglas-Rachford algorithm (Combettes and Pesquet, 2007) in the context of proximal splitting theory (Combettes and Pesquet, 2011). To set a meaningful bound K we have followed the criterion that at convergence the weighted-ℓ_1 norm of a matrix, as defined in section 2.2, mimics the ℓ_0 norm – as in formulation (3) –. K is then heuristically fixed as $K = 3N$, as it represents a conservative bound on the total number of fibre orientations to be identified, computed as the number of voxels N times an average bound on the number of fibre orientation per voxel. We initialise $\tau^{(0)}$ as the variance of the solution after the first iteration $\mathsf{X}^{(0)}$ and, in subsequent iterations, we update $\tau^{(t+1)} = \beta\tau^{(t)}$ with $\beta = 10^{-1}$. Ideally $\tau^{(t)}$ should decrease to 0 but we heuristically fix a lower bound $\tau_{\text{thr}} = 10^{-7}$, above which significant signal components could be identified. Experiments show that for a convergence bound $\nu = 10^{-3}$ the reweighting process stops after a relatively small number of iterations, typically 4 or 5. In our simulations, ν is set to 10^{-3} and $N_{\max}$ to 10.

To extract the final fibre directions from the solution to algorithm 1 in every voxel we perform a search for local maxima among all directions within a cone of 15° around every direction. In this entire process, we disregard the directions with contributions (i.e. coefficients) smaller than 10% of the maxima in order to filter out spurious peaks.

2.4. Phantom data

We perform our experiments using the phantom data used for the *HARDI reconstruction Challenge 2012* (Daducci et al., 2014a). The public results in (Daducci et al., 2014a) allow us to compare the performance of $\texttt{L2L0}_{\texttt{NW}}$ with other methods using different spatial regularisation schemes – such as TV regularisation mentioned above – with no need for their explicit implementation. The dataset is a $16 \times 16 \times 5$ volume that comprises 5 different fibre bundles that result in voxels with bending, crossing and kissing tracts. The response function of each bundle has been generated with a fractional anisotropy between 0.75 and 0.90 and the diffusion properties are constant along all its trajectory. More details on its geometry can be found in Daducci et al. (2014a).

The signal is contaminated with ***Rician noise*** (Gudbjartsson and Patz, 1995) as follows:

$$S_{\text{noisy}} = \sqrt{(S + \xi_1)^2 + (\xi_2)^2}, \tag{8}$$

with $\xi_1, \xi_2 \sim \mathcal{N}(0, \sigma^2)$ and $\sigma = S_0/\text{SNR}$ corresponding to a given signal-to-noise ratio (SNR) on the S_0 image. The quality of the reconstructions has been evaluated as a function of three different noise levels, i.e. $\text{SNR} = 10, 20, 30$ and 5 different q-space acquisition schemes (30, 20, 15, 10 and 6 samples), evenly spaced on half of the unit sphere.

2.5. Real Data

One HARDI dataset was acquired at $b = 3000$ s/mm^2 using 256 directions uniformly distributed on half of the unit sphere (as described by Jones et al. (1999)), TR/TE $= 7000/108$ ms and spatial resolution $= 2.5 \times 2.5 \times 2.5$ mm. To assess the robustness of $\texttt{L2L0}_{\texttt{NW}}$ to different under-sampling rates, the dataset has been retrospectively undersampled and three additional datasets have been created, consisting of only 30, 20 and 10 diffusion directions selected in order to be evenly spaced on half of the unit sphere using the tool *subsetpoints* which is available in the *camino* toolbox[2]. We will refer to these four data sets as $\texttt{hardi}_{256}$, $\texttt{hardi}_{30}$, $\texttt{hardi}_{20}$ and $\texttt{hardi}_{10}$, respectively. The actual SNR in the $b = 0$ images, computed as the ratio of the mean value in a region-of-interest placed in the WM and the standard deviation of the noise estimated in the background, was about 30.

[2]`www.camino.org.uk`

To evaluate the reconstructions from the undersampled real datasets, the metrics described in subsection 2.6 are computed considering the fully-sampled $\texttt{hardi}_{256}$ as the golden truth, as it is suggested by Yeh and Tseng (2013).

2.6. Evaluation criteria

To evaluate the quality of the reconstructions we have focussed on the performance of each method in both correctly assessing the number of fibre populations in each voxel and the angular accuracy in their orientation. In this work we adopted a set of metrics that Daducci et al. (2014a) used to evaluate and compare all methods participating in the *HARDI reconstruction Challenge 2012*. For consistency we have kept their notation to design the different quality indices. The *success rate* ($SR_\angle$) corresponds to the proportion of voxels in which a reconstruction algorithm correctly estimates the number of fibre populations. A fibre is considered to be correctly identified when an estimated fibre falls within a tolerance cone around a true fibre. To compare our results with different algorithms evaluated in (Daducci et al., 2014a), in this work the tolerance was set to $20°$. *False positive* and *negative* rates ($n^+_\angle$ and $n^-_\angle$, respectively) are an average over all voxels of the number of over-/underestimated fibre populations per voxel.

The *angular accuracy* is measured through the *mean angular error* $\bar{\theta}$ (in degrees) averaged over all true fibre directions, where the angular error associated with each true fibre is formally defined as:

$$\theta = \frac{180}{\pi} \arccos(|\boldsymbol{d}_{\text{true}} \cdot \boldsymbol{d}_{\text{estimated}}|), \tag{9}$$

where $\boldsymbol{d}_{\text{true}}$ and $\boldsymbol{d}_{\text{estimated}}$ are unitary vectors in the true fibre direction and the closest estimated direction. Note that indices $SR_\angle$, $n^+_\angle$ and $n^-_\angle$ represent mean values over all voxels of interest, whereas $\bar{\theta}$ is computed voxelwise and we study its statistical distribution to evaluate the general angular accuracy of each reconstruction.

2.7. Experimental setup

In the next section, we evaluate the quality of reconstructions using $\texttt{L2L0}_{\texttt{NW}}$, both for numerical simulations and tests on real data. Daducci et al. (2014b) showed that

L2L0 outperforms other state-of-the-art local methods that recover the FOD in the framework of spherical deconvolution. Consequently, we have chosen it as a benchmark to compare $\mathtt{L2L0}_{\mathtt{NW}}$ with respect to methods that perform voxel-by-voxel reconstruction of the fibre configuration. We had access to the original implementation by Daducci et al. (2014b) to run L2L0 reconstructions.

We also compare the performance of $\mathtt{L2L0}_{\mathtt{NW}}$, which jointly estimates the FOD and applies spatial regularisation, with respect to applying first a non-local denoising procedure and subsequently perform local reconstruction. We have chosen an adaptation of the Linear Minimum Mean Squared Error (LMMSE) filter proposed by Tristán-Vega and Aja-Fernández (2010) to simultaneously filter all different gradient images. We use a publicly available implementation of the Joint Anisotropic LMMSE filter[3] and subsequently apply L2L0 to reconstruct the FOD. We refer to this alternative as JAMMLSE+L2L0.

In addition, taking the advantage of the public results of the *HARDI reconstruction Challenge 2012* (Daducci et al., 2014a), we can compare the performance of $\mathtt{L2L0}_{\mathtt{NW}}$ with a representative collection of state-of-the-art methods for simulations on phantom data. In particular, we are able to establish a comparison with other methods using different spatial regularisation schemes – such as TV regularisation mentioned above – with no need for an explicit implementation of these methods.

Our optimisation code[4] was implemented in MATLAB and run on a standard 2.4 GHz Intel Xeon processor. The non-optimised version of the code is able to reconstruct a whole brain volume of $106 \times 106 \times 51$ voxels within approximately 4 hours.

3. Results and discussion

3.1. Phantom data

In this subsection we start comparing in detail the performance for $\mathtt{L2L0}_{\mathtt{NW}}$ relative to L2L0 and JAMMLSE+L2L0 for the phantom data set described in subsection 2.4. The performance of the three methods as a function of the undersampling

[3] http://www.nitrc.org/projects/jalmmse_dwi/
[4] Code is available at https://github.com/basp-group/co-dmri.

rate in q-space is reported in Figure 4. We consider 5 different acquisitions schemes (30, 20, 15, 10 and 6 samples) and present results for two different noise levels, at SNR $= 30$ and SNR $= 20$. The plots demonstrate that $\texttt{L2L0}_{\texttt{NW}}$ outperforms `L2L0` and `JAMMLSE+L2L0` for all number of samples, in both noise conditions. $\texttt{L2L0}_{\texttt{NW}}$ exhibits an accurate reconstruction ($SR_\angle \geq 85$ and $\mathrm{mean}(\bar{\theta}) \leq 6.5°$), robust to noise for different undersampling regimes, down to 15 samples. Denoising high-SNR data prior to reconstruction, as it is done in `JAMMLSE+L2L0`, seems not to improve the quality of the reconstructions. Indeed, at SNR $= 30, 20$ `JAMMLSE+L2L0` exhibits slightly worse results than `L2L0` (moderatly lower $SR_\angle$ and $\bar{\theta}$). With high quality data (SNR $= 30$ and from 30 to 15 samples), the differences between the three methods are fairly mild. The superiority of $\texttt{L2L0}_{\texttt{NW}}$ compared to `L2L0` and `JAMMLSE+L2L0` appears clearer as we move to higher undersampling regimes and SNR $= 20$, specially in terms of the ability of identifying the correct number of fibres (higher $SR_\angle$). The overall improvement in terms of the success rate is even more evident when we go down to 10 samples, where `L2L0` and `JAMMLSE+L2L0` exhibit a severe drop of the performance with $SR_\angle = 52$ (`L2L0`) and $SR_\angle = 50$ (`JAMMLSE+L2L0`) at SNR $= 30$ and $SR_\angle = 36$ (`L2L0`) and $SR_\angle = 38$ (`JAMMLSE+L2L0`) at SNR $= 20$, while $SR_\angle = 81$ (SNR $= 30$) and $SR_\angle = 72$ (SNR $= 20$) are obtained with $\texttt{L2L0}_{\texttt{NW}}$. We notice a significant deterioration of the reconstructions with all methods when decreasing the number of samples down to 6.

A more detailed analysis in severe noise conditions (SNR $= 10$) is presented in Figure 5. The plots show an important difference between the performance achieved by `L2L0`, that solves the problem voxelwise, and $\texttt{L2L0}_{\texttt{NW}}$ and `JAMMLSE+L2L0` that take into account the correlation between voxels and directions. At SNR=10, the denoising step in `JAMMLSE+L2L0`, specially indicated to correct the effect of the Rician noise at low SNR regimes (Tristán-Vega and Aja-Fernández, 2010), improves drastically the quality of the reconstructions. In particular, the overall $\bar{\theta}$ performances differ significantly between `L2L0` and `JAMMLSE+L2L0`, with an average enhancement of up to $5°$ in the mean $\bar{\theta}$ in different undersampling regimes. While in terms of angular resolution both $\texttt{L2L0}_{\texttt{NW}}$ and `JAMMLSE+L2L0` exhibit similar performance, $\texttt{L2L0}_{\texttt{NW}}$ shows a higher $SR_\angle$ down to 10 samples. In this noise setting, we analyse in detail the ability of correctly assessing the number of fibres through the false positives and negatives rates. Results show the effectiveness of the spatial regularisation applied both in JAMMLSE+L2L0 and $\mathrm{L2L0}_{\mathrm{NW}}$, specially in avoiding overestimated directions (extreme decrease of $n^+_\angle$) even if the number of missed fibres ($n^-_\angle$) is also significantly decreased.

Plots analogous to Figures 4 and 5 can be found in (Daducci et al., 2014a), where an exhaustive comparison of all methods participating in the *HARDI reconstruction Challenge 2012*[5] is presented. The performance of these algorithms is evaluated on the same phantom used in our simulations by computing the same quality metrics described in the present paper ($SR_\angle, \bar{\theta}, n^+$ and n^-). Figure 6 shows a comparison of the performance of $\mathrm{L2L0}_{\mathrm{NW}}$ run with *15 samples* with the following eight representative methods participating in the Challenge[6]: **(i)** $\mathrm{DTI}_{\mathrm{neigh}}$, classical DTI method enhanced using contextual information (Prckovska et al., 2010); **(ii)** L2−L1−DL, method using dictionary learning in the framework of ℓ_2-ℓ_1 reconstruction (Donoho, 2006); **(iii - iv)** L2−L1−TV and L2−L1−TGV, using the ℓ_2-ℓ_1 problem formulation and including spatial regularisation schemes based on total variation and total generalised variation, respectively (Mani et al., 2014); **(v - vi)** L2−L2 and NN−L2, based on ℓ_2 norm priors (Ramirez-Manzanares et al., 2007; Canales-Rodriguez et al., 2009); **(vii)** DOT, classical diffusion orientation transform (Ozarslan et al., 2006); **(viii)** $\mathrm{DSI}_{\mathrm{LR}}$, classical DSI enhanced using Lucy-Richardson deconvolution (Canales-Rodriguez et al., 2010). For a more detailed explanation of each reconstruction method, you can refer to (Daducci et al., 2014a). Direct quantitative comparisons with all these standard state-of-the-art algorithms is not straightforward from the results, since every method was tested using different sampling schemes (different number of samples and distribution of points). Yet, $\mathrm{L2L0}_{\mathrm{NW}}$ can be positioned in the overall picture. In Figure 6, participant methods are sorted by the number of samples used for the reconstruction, increasing from left to right. The actual number of samples is indicated on the plot for every method. In mild noise conditions (SNR $=$ 30), $\mathrm{L2L0}_{\mathrm{NW}}$ is able to correctly assess the num-

[5] http://hardi.epfl.ch/static/events/2012_ISBI

[6] For the sake of consistency, all methods are named following the same notation as in (Daducci et al., 2014a).

ber of fibres in 85% of voxels ($SR_\angle = 85$) using as few as 15 signal samples and this quality appears comparable to the best $SR_\angle$ scores obtained in the Challenge with methods using many more points (from 30 up to 257) to recover the fibre configuration. The superiority of $\texttt{L2L0}_{\texttt{NW}}$ appears to be even more significant when a more noisy setting is considered. At $\mathrm{SNR} = 10$, $\texttt{L2L0}_{\texttt{NW}}$ using only 15 samples, shows the same quality of reconstruction, in terms of both $SR_\angle$ and $\bar{\theta}$, as DSI using an exhaustive cartesian sampling scheme of 257 points. NN-L2 stands as the only method presenting slightly better results in terms of $SR_\angle$, yet, using 48 samples. Only with 15 samples $\texttt{L2L0}_{\texttt{NW}}$ is able to attain comparable levels of performance, thus implying a speed-up factor of three. We pay special attention to the comparison with the rest of methods that promote any kind of spatial regularisation. $\texttt{L2L0}_{\texttt{NW}}$ with 15 samples ($SR_\angle = 85$ and $\mathrm{mean}(\bar{\theta}) = 6.4°$) outperforms L2-L1-TV, the method imposing TV regularisation (Daducci et al. (2014a); see also Mani et al. (2014)), in terms of success rate ($SR_\angle = 75$) and present similar average angular error ($\mathrm{mean}(\bar{\theta}) = 6°$), stressing the fact that the latter uses a sampling scheme with the double number of points (30 samples). Overall, we point out that all participant methods imposing spatial regularisation (L2-L1-TV, L2-L1-TGV) use a significant amount of measurements (from 30 to 64 points) to recover the fibre configuration. The anatomical structured sparsity prior that we impose allows us to yield the same quality in the reconstructions using higher undersampling regimes.

3.2. Real Data

3.2.1. Quantitative comparison

In this subsection, we compare quantitatively the reconstructions obtained from undersampled real data (i.e. $\texttt{hardi}_{30}$, $\texttt{hardi}_{20}$ and $\texttt{hardi}_{10}$) to those with fully-sampled data (i.e. $\texttt{hardi}_{256}$), considering the latter as ground-truth, for L2L0, JAMMLSE+L2L0 and $\texttt{L2L0}_{\texttt{NW}}$. Results quoted next are in agreement with those obtained for numerical simulations on the phantom, confirming that $\texttt{L2L0}_{\texttt{NW}}$ actually outperforms L2L0 and JAMMLSE+L2L0. Bearing in mind that he actual SNR in the $b = 0$ images is about 30, results for JAMMLSE+L2L0 and $\texttt{L2L0}_{\texttt{NW}}$ appear in line with conclusions driven from the *HARDI Reconstruction Challenge 2012*, where it was

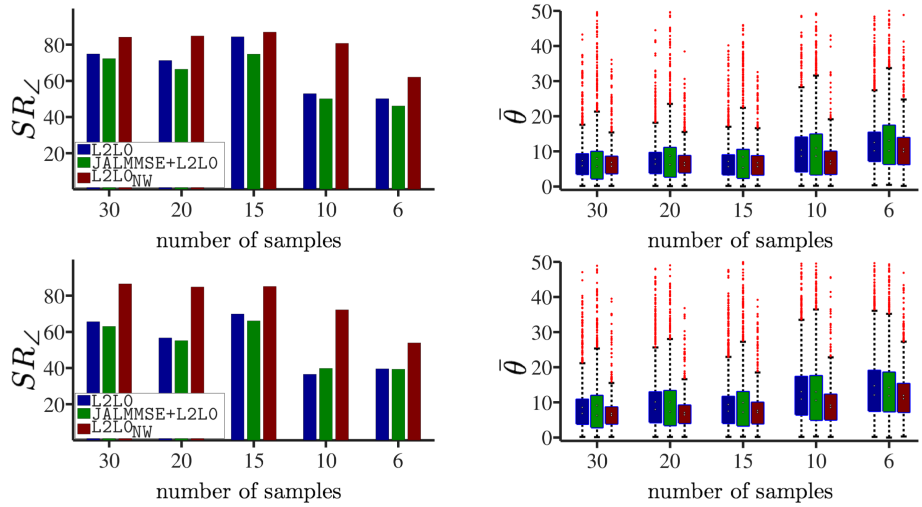

Figure 4: Comparison of $SR_\angle$ and $\bar{\theta}$ between L2L0, JAMMLSE+L2L0 and L2L0$_{\mathrm{NW}}$ approaches. Experiments are performed on the phantom dataset used in Daducci et al. (2014a) for a fixed SNR $= 30$ (top row) and SNR $= 20$ (bottom row). On the left, $SR_\angle$ represents the *success rate*. On the right, the boxplot diagrams present the distribution of $\bar{\theta}$, with the edges of each box representing the 25th and 75th percentiles, the mean and median value appear as "square" and "circle" value and the outliers are plotted as red dots.

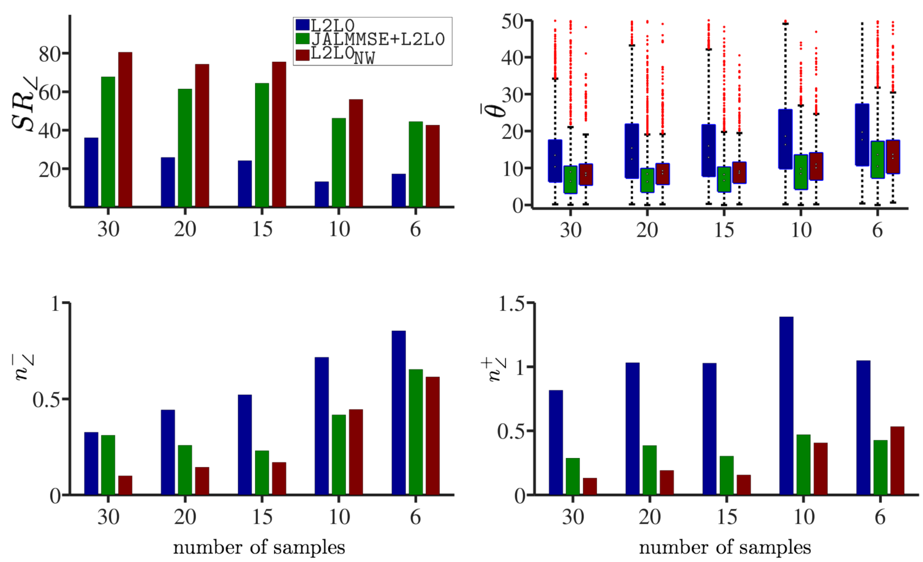

Figure 5: Comparison of $SR_\angle$ and $\bar{\theta}$ between L2L0, JAMMLSE+L2L0 and L2L0$_{\mathrm{NW}}$ approaches. Experiments are performed on the phantom dataset used in Daducci et al. (2014a) for a fixed SNR $= 10$. On the top left, $SR_\angle$ represents the *success rate*. On the top right, the boxplot diagrams present the distribution of $\bar{\theta}$ with the same conventions as for Figure 4. On the bottom row, $n_\angle^-$ and $n_\angle^+$ represent the false negatives and positives rates.

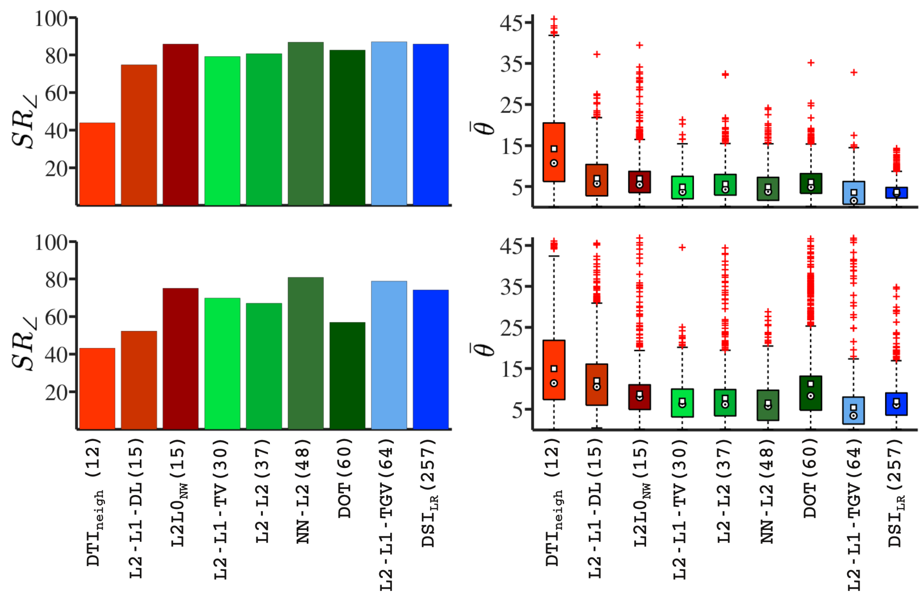

Figure 6: Comparison of $SR_\angle$ and $\bar{\theta}$ between different reconstruction methods. Experiments are performed on the phantom dataset used in Daducci et al. (2014a) for a fixed $\mathrm{SNR} = 30$ (top row) and $\mathrm{SNR} = 10$ (bottom row). On the left, $SR_\angle$ represents the *success rate*. For the sake of comparison, the number of samples used for the reconstruction is reported in parentheses next to the name of each method. On the right, the boxplot diagrams present the distribution of $\bar{\theta}$, with the edges of each box representing the 25th and 75th percentiles, the mean and median value appear as "square" and "circle" value and the outliers are plotted as red dots.

shown that spatial regularisation appeared to be effective also in low noise regimes, while merely denoising the images did not (Daducci et al., 2014a).

The average *mean angular error* ($\bar{\theta}$) using 30 samples was $13.9° \pm 11.4°$ (mean $\pm$ standard deviation over WM voxels of the whole brain volume) for L2L0, $14.5° \pm 10.8°$ for JAMMLSE+L2L0 and $7.8° \pm 9.14°$ for L2L0$_{\mathrm{NW}}$. Reconstructions using 20 samples had an average error of $15.7° \pm 11.2°$ for L2L0, $16.7° \pm 11.8°$ for JAMMLSE+L2L0 and $9.1° \pm 9.6°$ for L2L0$_{\mathrm{NW}}$. When one goes down to 10 samples, reconstructions using L2L0 and JAMMLSE+L2L0 exhibit an angular error of $19.8° \pm 11.25°$ and $19.8° \pm 12.0°$, respectively, which is already higher than the resolution of the spherical discretisation defined by our dictionary; while the angular error for L2L0$_{\mathrm{NW}}$ is $13.6° \pm 10.5°$. Results for the success rate are as well consistent with the results obtained in simulations. As in numerical simulations, the benefits of imposing a spatial regularisation directly on the fibre orientations are more remarkable when we go to higher subsampling regimes. The $SR_{\angle}$ was $31.1\% \pm 46.3\%$ for L2L0, $34.8\% \pm 47.6\%$ for JAMMLSE+L2L0 and $67.0\% \pm 47.0\%$ for L2L0$_{\mathrm{NW}}$ with 30 samples; $27.9\% \pm 44.9\%$ for L2L0, $28.0\% \pm 45.0\%$ for JAMMLSE+L2L0 and $61.7\% \pm 48.6\%$ for L2L0$_{\mathrm{NW}}$ at 20 samples. All methods present a degradation in the quality of their reconstructions when we go down to 10 samples, $SR_{\angle}$ decreasing to $16\% \pm 36.6\%$ for L2L0, $18.8\% \pm 39.0\%$ for JAMMLSE+L2L0 and $40.6\% \pm 49.1\%$ for L2L0$_{\mathrm{NW}}$.

Figures 7 and 8 illustrate the numerical results for one representative slice of the brain volume. The angular accuracy of each reconstruction is presented by plotting the mean angular error $\bar{\theta}$ per voxel in Figure 7. A map of the *number of false positives* and *false negatives* per voxel is used to illustrate the ability of each method of correctly assessing the number of fibres in Figure 8. The images show the superiority of L2L0$_{\mathrm{NW}}$ with respect to L2L0 and JAMMLSE+L2L0, specially in those voxels close to the boundaries with the grey matter and the cerebrospinal fluid.

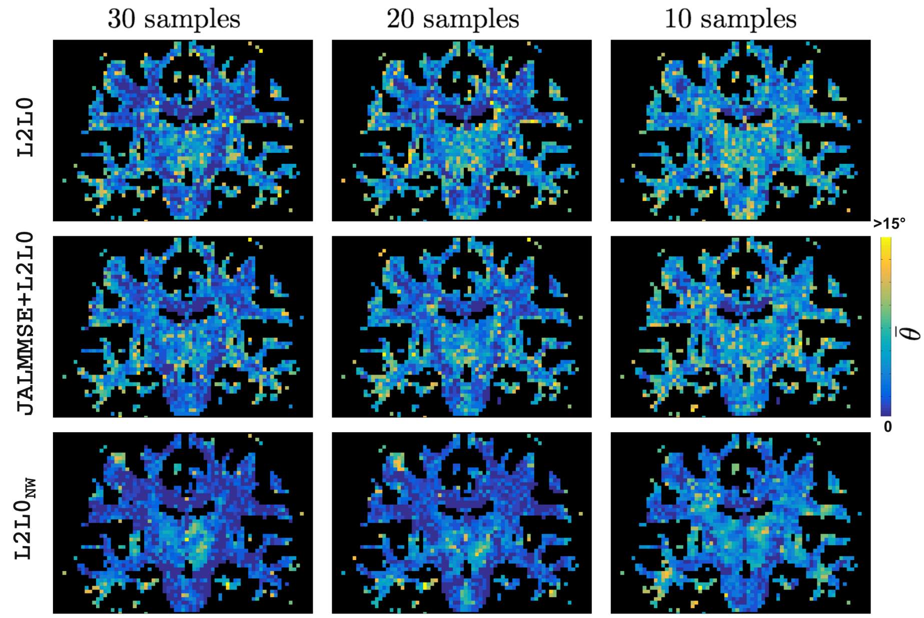

Figure 7: **Angular accuracy** (map of $\bar{\theta}$ per voxel) in real data between `L2L0`, `JAMMLSE+L2L0` and $\texttt{L2L0}_{\texttt{NW}}$ reconstructions with 30, 20 and 10 samples ($\texttt{hardi}_{30}$, $\texttt{hardi}_{20}$, $\texttt{hardi}_{10}$ datasets, respectively).

3.2.2. Qualitative comparison

The reconstructions[7] of the FOD obtained with `L2L0` and $\texttt{L2L0}_{\texttt{NW}}$ for a significant slice of the brain in the corona radiata region are compared qualitatively in Figures 9 and 10. These plots show the robustness of each method to two different undersampling regimes, $\texttt{hardi}_{30}$ and $\texttt{hardi}_{10}$. In the light of the quantitative results obtained for both phantom and real data and given the fact that qualitative differences between reconstructions using `L2L0` and `JAMMLSE+L2L0` are difficult to appreciate, we do not show qualitative results for `JAMMLSE+L2L0`. In all images, three meaningful regions with fibre bundle crossings have been highlighted. With 30 samples (Figure 9 corresponding to $\texttt{hardi}_{30}$), the FODs reconstructed by $\texttt{L2L0}_{\texttt{NW}}$ present neater and sharper profiles with less presence of spurious peaks than the ones reconstructed by `L2L0`. In addition, the fibre orientation distribution field reconstructed by $\texttt{L2L0}_{\texttt{NW}}$

[7]The images have been created using the tool *mrview* of *mrtrix*. This required the FOD from `L2L0` and $\texttt{L2L0}_{\texttt{NW}}$ to be previously converted to spherical harmonics.

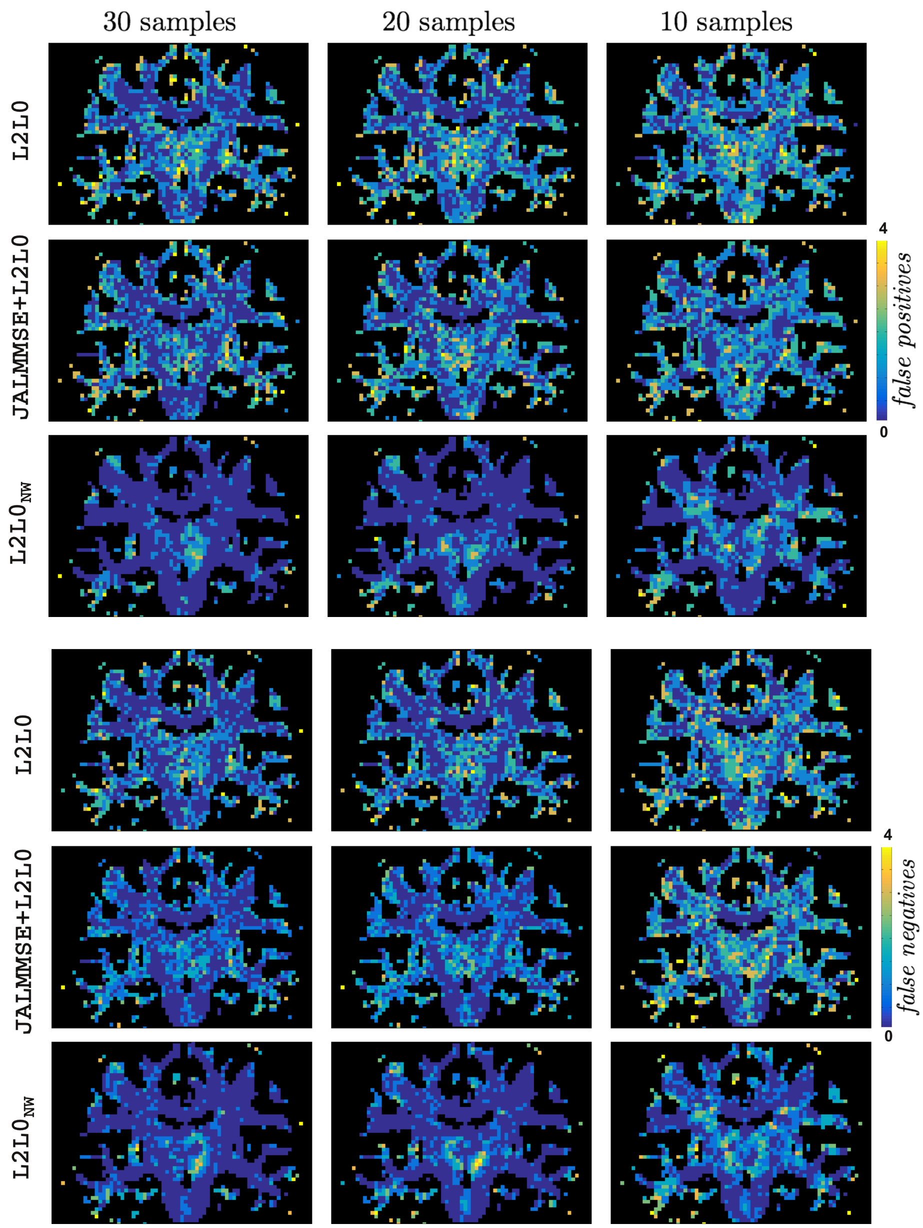

Figure 8: **Ability of correctly assessing the number of fibres** in real data between `L2L0`, `JAMMLSE+L2L0` and $\texttt{L2L0}_{\texttt{NW}}$ reconstructions with 30, 20 and 10 samples ($\texttt{hardi}_{30}$, $\texttt{hardi}_{20}$, $\texttt{hardi}_{10}$ datasets, respectively). Map of number of *false positives* (top) and *false negatives* (bottom) per voxel.

looks qualitatively smoother overall. As a consequence, fibre bundles are better defined through more clearly identified peaks. Plots in Figure 10 show reconstructions performed with only 4% of the original data (10 samples). In these images – corresponding to reconstructions with highly undersampled data – the above-mentioned qualitative differences between the two methods are confirmed and even more easily noticeable. As discussed in section 1, these differences can have a significant impact when applying tractography methods on these fibre orientation fields.

4. Discussion and conclusions

In this work we have proposed a novel algorithm to recover the intra-voxel FOD simultaneously for all voxels. The method leverages a spatially structured sparsity prior directly on the FOD, where the structure originates from the spatial coherence of the fibre orientation between neighbour voxels. We have made use of a reweighting scheme to enforce structured sparsity in the solution. We have shown through numerical simulations and tests on real data that this method outperforms a voxel-by-voxel reconstruction method when assessing the correct number of fibres and the angular precision of their orientation. As shown in section 3, exploiting spatial information about the neighbouring directions appears essential to ensure a stronger robustness to noise and ability to go to higher undersampling regimes, leading to accurate reconstructions with only 15 samples.

We also compare the performance of our proposed method with respect to applying first a non-local denoising procedure and subsequently perform local reconstruction. This comparison allows us to highlight the benefits of using a spatial regularisation as in our approach as opposed to this decoupled strategy. As presented in simulations, our spatial prior on the FOD outperforms as well the empirical TV regularisation of q-space images proposed by Mani et al. (2014), being able to recover the fibre orientation distribution using fewer samples. Note that spatial regularisation of the q-space images is actually complementary to our formulation and could be added as an additional prior to our method.

The regularisation presented in this paper could as well be applied in a voxel-by-

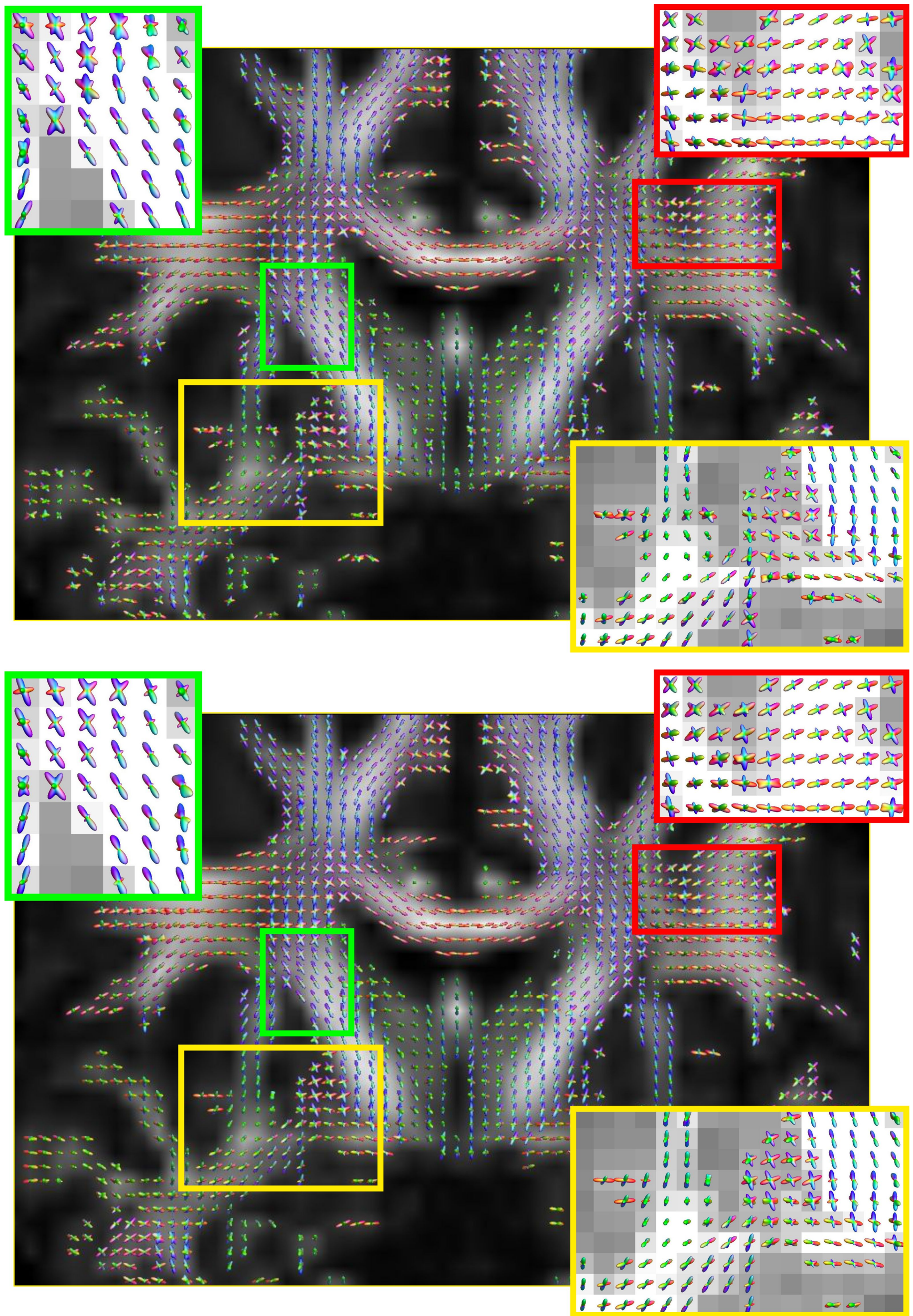

Figure 9: Qualitative comparison on HARDI human data. Reconstructions of the FODs in the corona radiata region are shown for `L2L0` (top) and $\texttt{L2L0}_{\texttt{NW}}$ (bottom) for 30 samples superimposed to the FA map.

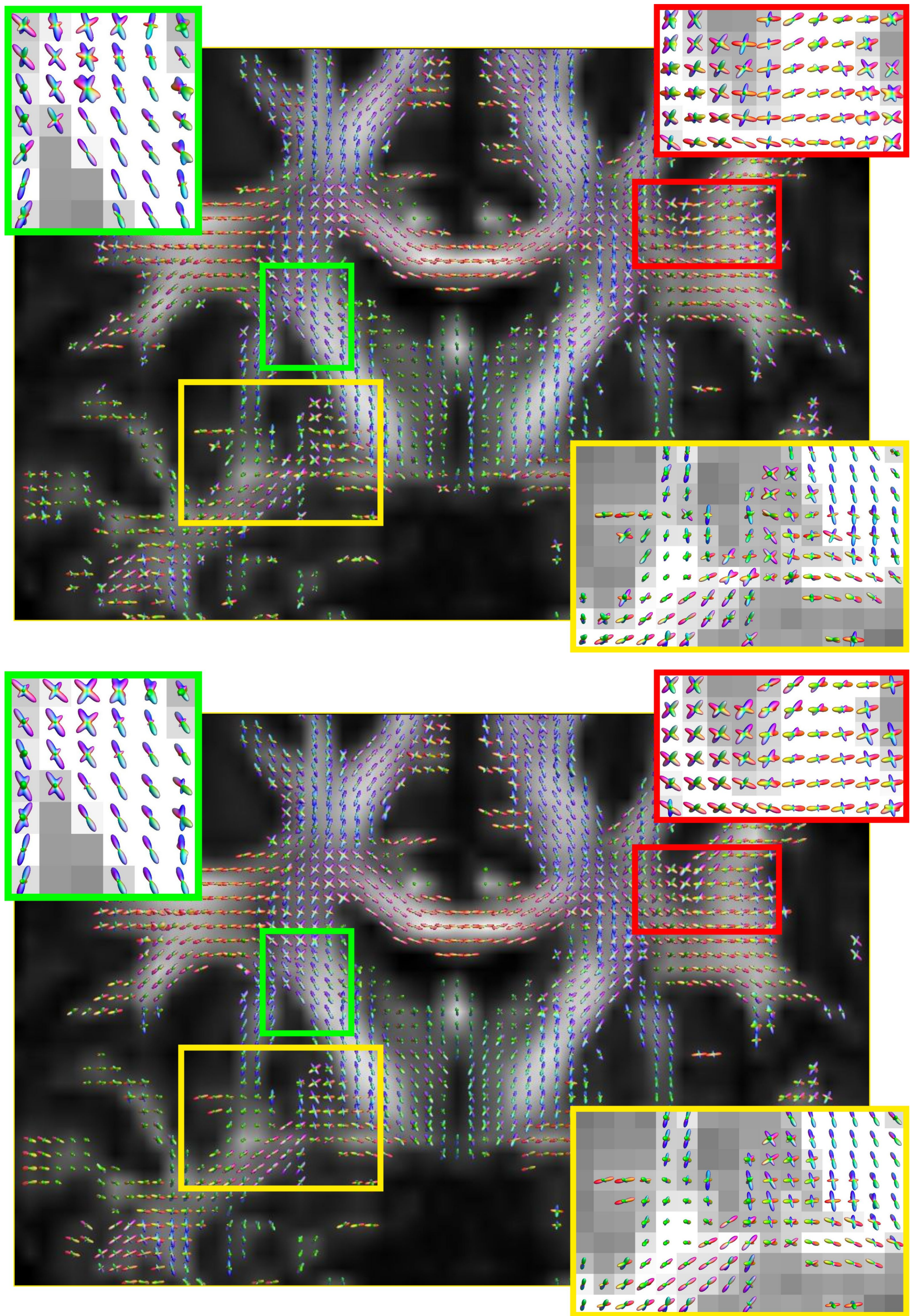

Figure 10: Qualitative comparison on HARDI human data. Reconstructions of the FODs in the corona radiata region are shown for L2L0 (top) and $\mathtt{L2L0}_{\mathtt{NW}}$ (bottom) for 10 samples superimposed to the FA map.

voxel configuration, redefining the weights in formulation (3) to account for the values of the FOD in a defined neighbourhood. Preliminary investigations in this direction did not provide promising results. Fixing a single bound to estimate the number of fibres *separately* in every voxel of the brain appears to be too constraining. On the contrary, setting a bound on the total number of fibres of the whole volume and solving the problem for all voxels simultaneously leaves more freedom on the effective directions (number of non-zero coefficients) per voxel. Furthermore, future evolutions of this algorithm should enable undersampling in Fourier space (k-space) for each of the q-space images acquired. This combined $k-q$-space sampling approach, along the lines of work by Mani et al. (2014), will potentially enable a significant additional acceleration, in which context a voxel-by-voxel approach is not an option. Regarding computing resources, the memory requirements of a reweighting scheme to solve each voxel independently but using neighbourhood information to define the weights would not differ from $\texttt{L2L0}_{\texttt{NW}}$, bearing in mind that the main operator Φ remains exactly the same for both formulations (3) and (5). In any case, the computation time of $\texttt{L2L0}_{\texttt{NW}}$ is affordable for a single processor, as described in section 2.7.

In recent work, Daducci et al. (2015) present a general framework for Accelerated Microstructure Imaging via Convex Optimization (AMICO) to recover the microstructure configuration voxel-by-voxel in regions with one single fibre population. Future investigations will consider to exploit the spatial coherence of the microstructural features of the fibres all over the brain with the aim of extending the AMICO framework to regions of the WM with multiple fibre populations and more complex configurations.

Acknowledgements

The authors thank R. Carrillo for insightful discussions on optimisation algorithms and constructive comments on the manuscript. A. Auría is supported by the Swiss National Science Foundation (SNSF) under grant 205321-138311. This work is supported by the Center for Biomedical Imaging (CIBM) of the Geneva-Lausanne Universities and the EPFL, as well as the foundations Leenaards and Louis-Jeantet.

References

Alexander, D., 2005. Maximum entropy spherical deconvolution for diffusion MRI. Information Processing in Medical Imaging (IPMI) , 76–87.

Basser, P., Mattiello, J., Le Bihan, D., 1994. MR diffusion tensor spectroscopy and imaging. Biophysical journal 66, 259–267.

Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system a technical review. NMR in Biomedicine 15, 435–455.

Canales-Rodriguez, E., Iturria-Medina, Y., Aleman-Gomez, Y., Melie-Garcia, L., 2010. Deconvolution in diffusion spectrum imaging. NeuroImage 50, 136–149.

Canales-Rodriguez, E., Lin, C.P., Iturria-Medina, Y., Yeh, C.H., Cho, K.H., Melie-Garcia, L., 2009. Diffusion orientation transform revisited. NeuroImage 49, 1326–1339.

Candès, E., Romberg, J., Tao, T., 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory 54, 489–509.

Candès, E., Wakin, M., Boyd, S., 2008. Enhancing sparsity by reweighted ℓ_1 minimization. Journal of Fourier Analysis and Applications 14, 877–905.

Combettes, P.L., Pesquet, J.C., 2007. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. IEEE Journal of Selected Topics in Signal Processing 1, 564–574.

Combettes, P.L., Pesquet, J.C., 2011. Proximal splitting methods in signal processing. Fixed-Point Algorithms for Inverse Problems in Science and Engineering (Bauschke H. H., Burachik R.S., Combettes P.L., Elser V., Luke D.R., Wolkowicz H., eds) , 185–212.

Daducci, A., Canales-Rodriguez, E., Descoteaux, M., Garyfallidis, E., Gur, Y., Lin, Y., Mani, M., Merlet, S., Paquette, M., Ramirez-Manzanares, A., Reisert, M., Rodrigues, P., Sepehrband, F., Caruyer, E., Choupan, J., Deriche, R., Jacob, M., Menegaz, G., Prckovska, V., Rivera, M., Wiaux, Y., Thiran, J., 2014a. Quantitative comparison of reconstruction methods for intra-voxel fiber recovery from diffusion mri. IEEE Transactions on Medical Imaging 33, 384–399.

Daducci, A., Canales-Rodriguez, E.J., Zhang, H., Dyrby, T.B., Alexander, D.C., Thiran, J.P., 2015. Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. Neuroimage 105, 32–44. URL: https://github.com/daducci/AMICO, doi:10.1016/j.neuroimage.2014.10.026.

Daducci, A., Ville, D.V.D., Thiran, J., Wiaux, Y., 2014b. Sparse regularization for fiber ODF reconstruction: from the suboptimality of ℓ_2 and ℓ_1 priors to ℓ_0. Medical Image Analysis 18, 820–33.

Donoho, D., 2006. Compressed sensing. IEEE Transactions on Information Theory 54, 1289–1306.

Fillard, P., Pennec, X., Arsigny, V., Ayache, N., 2007. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. IEEE Transactions on Medical Imaging 26, 1472–1482. URL: http://www.inria.fr/sophia/asclepios/Publications/Pierre.Fillard/Fillard.TMI.2007.pdf, doi:10.1109/TMI.2007.899173.

Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy MRI data. Magnetic Resonance in Medicine 34, 910–914.

Huang, J., Yagel, R., Filippov, V., Kurzion, Y., 1998. An accurate method for voxelizing polygon meshes. Proceedings of the 1998 IEEE Symposium on Volume Visualization , 119–126.

Jian, B., Vermuri, B., 2007. A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. IEEE Transactions on Medical Imaging 26, 1464–1471.

Jones, D., Horsfield, M., Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magnetic Resonance in Medicine 42, 515–525.

Mani, M., Jacob, M., Guidon, A., Magnotta, V., Zhong, J., 2014. Acceleration of high angular and spatial resolution diffusion imaging using compressed sensing with multichannel spiral data. Magnetic Resonance in Medicine .

Michailovich, O., Rathi, Y., Dolui, S., 2011. Spatially regularized compressed sensing for high angular resolution diffusion imaging. IEEE Transactions on Medical Imaging 30, 1100–1115.

Nocedal, J., Wright, S., 2006. Numerical Optimization. Springer.

Ozarslan, E., Shepherd, T., Vemuri, B., Blackband, S., Mareci, T., 2006. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). NeuroImage 31, 1086–1103.

Prckovska, V., Rodrigues, P., Duits, R., ter Haar Romeny, B., Vilanova, A., 2010. Extrapolating fiber crossings from DTI data. Can we gain similar information as HARDI? Proc. Workshop computational diffusion MRI, MICCAI , 26–37.

Ramirez-Manzanares, A., Rivera, M., Vemuri, B., Carney, P., Mareci, T., 2007. Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry. IEEE Transactions on Medical Imaging 26, 1091–1102.

Rudin, L.I., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. Phys. D 60, 259–268. URL: `http://dx.doi.org/10.1016/0167-2789(92)90242-F`, doi:`10.1016/0167-2789(92)90242-F`.

Tournier, J., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. NeuroImage 35, 1459–1472.

Tournier, J., Calamante, F., Gadian, D., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage 23, 1176–1185.

Tristán-Vega, A., Aja-Fernández, S., 2010. Dwi filtering using joint information for dti and hardi. Medical Image Analysis 14, 205–218.

Wedeen, V.J., Hagmann, P., Tseng, W., Reese, T., Weisskoff, R., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic ressonance imaging. Magnetic Ressonance in Medicine 54, 1377–1386.

Wiest-Daessl, N., Prima, S., Coup, P., Morrissey, S.P., Barillot, C., 2008. Rician noise removal by non-local means filtering for low signal-to-noise ratio mri: Applications to dt-mri., Springer. pp. 171–179.

Yeh, F., Tseng, W., 2013. Sparse solution of fiber orientation distribution function by diffusion decomposition. PLoS One 8.