

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/162488/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Powell, Elizabeth, Battocchio, Matteo, Parker, Christopher S. and Slator, Paddy J. 2021. Generalised hierarchical bayesian microstructure modelling for diffusion MRI. Presented at: International Workshop on Computational Diffusion MRI, 01 October 2021. Published in: Cetin-Karayumak, S. ed. Computational Diffusion MRI. CDMRI 2021. Lecture Notes in Computer Science , vol.13006 Springer Nature, 36–47. 10.1007/978-3-030-87615-9_4

Publishers page: https://doi.org/10.1007/978-3-030-87615-9_4

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Generalised Hierarchical Bayesian Microstructure Modelling for Diffusion MRI

Elizabeth Powell¹, Matteo Battocchio², Christopher S. Parker¹, and Paddy J. Slator^{1*}

¹ Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK

² Department of Computer Science, University of Verona, Verona, Italy
*p.slator@ucl.ac.uk

Abstract. Microstructure imaging combines tailored diffusion MRI acquisition protocols with a mathematical model to give insights into sub-voxel tissue features. The model is typically fit voxel-by-voxel to the MRI image with least squares minimisation to give voxelwise maps of parameters relating to microstructural features, such as diffusivities and tissue compartment fractions. However, this fitting approach is susceptible to voxelwise noise, which can lead to erroneous values in parameter maps. Data-driven Bayesian hierarchical modelling defines prior distributions on parameters and learns them from the data, and can hence reduce such noise effects. Bayesian hierarchical modelling has been demonstrated for microstructure imaging with diffusion MRI, but only for a few, relatively simple, models. In this paper, we generalise hierarchical Bayesian modelling to a wide range of multi-compartment microstructural models, and fit the models with a Markov chain Monte Carlo (MCMC) algorithm. We implement our method by utilising Dmipy, a microstructure modelling software package for diffusion MRI data. Our code is available at github.com/PaddySlator/dmipy-bayesian.

Keywords: Bayesian statistics · Bayesian hierarchical model · Microstructure modelling · Diffusion MRI

1 Introduction

Diffusion MRI (dMRI) measures the microscopic motion of water molecules, and is hence sensitive to tissue microstructure. Microstructural modelling combines specifically-designed dMRI acquisitions with a tissue model to enable estimation of parameters relating to tissue microstructure. These techniques have been widely applied in neuroimaging, with prominent examples of brain microstructure imaging including NODDI [10], the standard model of diffusion in neuronal tissue [7] and the spherical mean technique [4]. Microstructural modelling has also provided insights into body MRI [5], for example in prostate cancer [9].

The core fitting procedure in microstructure imaging estimates model parameters given the observed dMRI signal (Figure 1, top panel). The vast majority of

fitting techniques assume that voxels are independent; in other words, the model is separately fit to the signal in each voxel, usually with nonlinear least squares estimation. An alternative approach is to use an MCMC algorithm to estimate parameter posterior distributions in each voxel [3]. Orton et al. [8] introduced a hierarchical Bayesian model fitting approach for the intravoxel incoherent motion (IVIM [6]) model. Their model breaks the assumption of independent pixels by introducing a Gaussian prior (estimated from the data) over the microstructural model parameters across a region of interest (ROI). By using an MCMC algorithm to fit the Bayesian model, they showed an improvement in IVIM parameter maps of the liver. This approach has also been applied to combined T₂-IVIM modelling in the placenta [2].

In this paper, we generalise the Bayesian approach to apply to any microstructural model, derive the corresponding MCMC algorithm, and implement arbitrary upper and lower parameter bounds. We also utilise regional priors, which may be more appropriate than a global prior for fitting across distinct neurological tissue types. The MCMC algorithm is implemented in Python by utilising the Diffusion Microstructure Imaging in Python (Dmipy [1]) software package. We demonstrate our algorithm on simulations and on HCP data, and show clear advantages over the standard least squares fitting technique.

2 Methods

2.1 General Bayesian Microstructure model

We extend the approach of Orton et al. [8] to a general multi-compartment microstructural model. A schematic of the hierarchical Bayesian framework is shown in Figure 1.

We consider a general multi-compartment model of N_{comp} compartments, with a set of underlying microstructure-related parameters θ . For notational convenience we group θ by parameter type as

$$\theta = \left\{ \{f_k\}_{k=1}^{N_{comp}-1}, \{x_j\}_{j=1}^J \right\} \quad (1)$$

where f_k denotes compartment signal fractions and x_j the other parameters, e.g. diffusivities, orientations, radii, etc. Assuming that relaxation times are fixed across compartments, the signal fractions sum to 1, i.e. $\sum_{k=1}^{N_{comp}} f_k = 1$, meaning that $f_{N_{comp}}$ is not a free parameter but fixed as $1 - \sum_{k=1}^{N_{comp}-1} f_k$.

A general microstructural model comprises a mapping - or *signal equation* - between underlying tissue-related parameters θ and acquisition parameters t_n (typically b-value and gradient direction), and a dMRI signal intensity S_n , i.e.

$$S_n = S_0 g_n(\theta, t_n) \quad (2)$$

where S_0 is the signal intensity without diffusion weighting. The experimentally-measured signal in the presence of noise is hence modelled as

$$y_n = S_n + \epsilon_n = S_0 g_n(\theta, t_n) + \epsilon_n \quad (3)$$

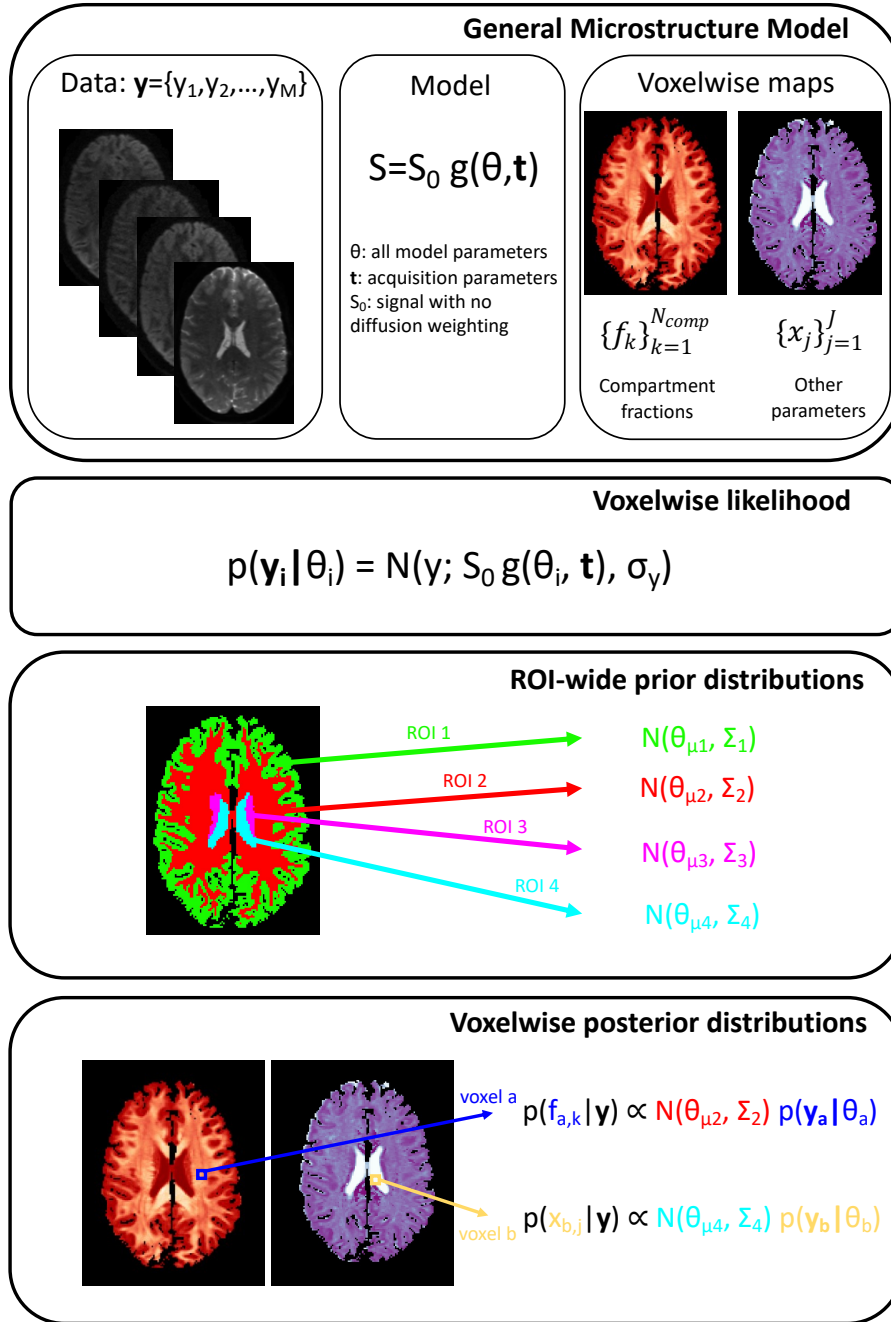


Fig. 1. Schematic of Bayesian hierarchical model for a general microstructure model. Top panel defines a general microstructure model, g that maps microstructure-related parameters θ and acquisition parameters \mathbf{t} to dMRI signal S . The microstructure parameters can be grouped by parameter type as $\theta = \left\{ \{f_k\}_{k=1}^{N_{comp}-1}, \{x_j\}_{j=1}^J \right\}$ where f_k are the compartment signal fractions and x_j are the other parameters. Second panel defines the voxelwise likelihood function. Third panel displays the ROI-wide Gaussian priors, note that θ_μ and Σ are learnt from the data for all ROIs. Fourth panel displays the voxelwise parameter posterior distributions and corresponding parameter maps.

where y_n is the measured signal and ϵ_n is noise.

We now consider all measurements for a voxel i - i.e. the signal intensities at all acquisition parameters $\mathbf{t} = \{t_1, \dots, t_N\}$ - which we denote $\mathbf{y}_i = [y_1, \dots, y_N]^T$. The likelihood, assuming normally distributed noise with variance σ_y^2 , is therefore

$$p(\mathbf{y}_i|\theta_i, S_0, \sigma_y^2) = (2\pi\sigma_y^2)^{-N/2} \exp\left(\frac{-1}{2\sigma_y^2} \sum_{n=1}^N (y_n - S_0 g_n(\theta_i, t_n))^2\right) \quad (4)$$

where θ_i denotes the microstructural model parameter values in voxel i . Orton et al. [8] demonstrated that the “nuisance parameters” S_0 and σ_y^2 can be marginalised out from Equation (4) to give the following marginalised likelihood

$$p(\mathbf{y}_i|\theta_i) \propto [\mathbf{y}_i^T \mathbf{y}_i - (\mathbf{y}_i^T \mathbf{g}_i)^2 / \mathbf{g}_i^T \mathbf{g}_i]^{-N/2} \quad (5)$$

where $\mathbf{g}_i = [g_1(\theta_i, t_1), \dots, g_N(\theta_i, t_N)]$ are the model predicted signals for voxel i .

2.2 Parameter Transforms

Microstructure model fitting needs to enforce physically reasonable minimum and maximum values of parameters; for example, diffusivities need to be positive. Here we generalise the transforms used by Orton et al. [8] to enable arbitrary minimum and maximum constraints. For a parameter p , we define a transform

$$p' = \log(p - p_{min}) - \log(p_{max} - p). \quad (6)$$

which maps the interval (p_{min}, p_{max}) to \mathbb{R} . By defining the Bayesian prior on the transformed parameter p' , we therefore constrain p between p_{min} and p_{max} . Default values for p_{min} and p_{max} are set as the minimum and maximum values defined in the Dmipy variable “model.parameter_ranges”; however, they can also be manually defined by the user.

2.3 Bayesian shrinkage priors

Orton et al. [8] used a multivariate Gaussian Bayesian shrinkage prior on the IVIM model parameters, with the prior defined over a single user-defined ROI. The Bayesian fitting method is generalised here to the multiple ROI case simply by running the derived MCMC algorithm separately on the voxelwise dMRI data from each ROI; however, note that here and throughout the methods section we consider the single ROI case for brevity, without loss of generality. The prior generalised for any microstructural model is denoted

$$p(\theta|\theta_\mu, \Sigma) = N(\theta; \theta_\mu, \Sigma) \quad (7)$$

where θ_μ is a vector whose elements encode the prior means of the parameters, Σ is their covariance and $N(\theta; \theta_\mu, \Sigma)$ denotes the multivariate normal probability density function (PDF) with variable θ , mean θ_μ and covariance Σ . Again, we emphasise that θ_μ and Σ are estimated from the data.

To generalise from Orton et al.'s [8] two-compartment model to an arbitrary multi-compartment model, all signal fractions must to sum to one. We enforce this (following Harms et al. [3]) by modifying the prior to

$$p(\theta_i|\theta_\mu, \Sigma) = \begin{cases} N(\theta_i; \theta_\mu, \Sigma) & \text{if } \sum_{j=1}^{n-1} f_j \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

To complete the model we define a hyper-prior on θ_μ and Σ as a non-informative Jeffrey's prior

$$p(\theta_\mu, \Sigma) = |\Sigma|^{-1/2} \quad (9)$$

2.4 Posterior Distributions

Each ROI has its own posterior distribution, which can be written as [8]

$$p(\theta_{1:M}, \theta_\mu, \Sigma | \mathbf{y}_{1:M}) \propto p(\mathbf{y}_{1:M} | \theta_{1:M}) p(\theta_{1:M} | \theta_\mu, \Sigma) p(\theta_\mu, \Sigma)$$

where $\theta_{1:M} = \{\theta_1, \theta_2, \dots, \theta_M\}$ are the parameters and $\mathbf{y}_{1:M} = \{\mathbf{y}_1, \dots, \mathbf{y}_M\}$ the dMRI data for all voxels in the ROI. Substituting in equations (5), (7), (9) gives

$$p(\theta_{1:M}, \theta_\mu, \Sigma | \mathbf{y}_{1:M}) \propto \left(\prod_{i=1}^M [\mathbf{y}_i^T \mathbf{y}_i - (\mathbf{y}_i^T \mathbf{g}_i)^2 / \mathbf{g}_i^T \mathbf{g}_i]^{-N/2} \right) \left(\prod_{i=1}^M N(\theta_i; \theta_\mu, \Sigma) \right) |\Sigma|^{-1/2}$$

from which we can draw samples with an MCMC algorithm.

2.5 MCMC algorithm

The MCMC algorithm is derived here, and given as pseudocode in Algorithm 1.

ROI-wide parameters Following Orton et al. [8], the MCMC updates for the ROI-wide prior parameters θ_μ and Σ are Gibbs moves. The conditional distributions are (up to proportionality)

$$p(\theta_\mu | \theta_{1:M}, \Sigma, \mathbf{y}_{1:M}) \propto \prod_{i=1}^M N(\theta_i; \theta_\mu, \Sigma) = N(\theta_\mu; m, V)$$

where $m = M^{-1} \sum_{i=1}^M \theta_i$ $V = M^{-1} \Sigma$, and the second line comes from rearranging the multivariate normal PDF so that θ_μ is the variable. The MCMC update is therefore sampled as follows

$$\theta_\mu \sim N(m, V) \quad (10)$$

where $N(m, V)$ is a multivariate normal distribution with mean m and covariance V . Following the same steps for Σ (see Orton et al. [8] for full details) gives the MCMC update for Σ

$$\Sigma \sim W^{-1}(\Phi, M - 3) \quad (11)$$

where $\Phi = \sum_{i=1}^M (\theta_i - \theta_{mu}) (\theta_i - \theta_{mu})^T$ and W^{-1} is the inverse-Wishart distribution.

Voxelwise parameters For the non-signal fraction voxelwise parameters the posterior distribution up to proportionality is

$$p(x_{i,j}|x_{i,-j}, \{f_{i,k}\}_{k=1}^{N_{comp}-1}, \theta_\mu, \Sigma) \propto [\mathbf{y}_i^T \mathbf{y}_i - (\mathbf{y}_i^T \mathbf{g}_i)^2 / \mathbf{g}_i^T \mathbf{g}_i]^{-N/2} N(\theta_i; \theta_\mu, \Sigma) \quad (12)$$

where $x_{i,j}$ is the value of parameter x_j in voxel i , $x_{i,-j} = \{x_{i,1}, \dots, x_{i,j-1}, x_{i,j+1}, \dots, x_{i,J}\}$ denotes the set of all non-signal fraction parameters except $x_{i,j}$, and $\{f_{i,k}\}_{k=1}^{N_{comp}-1}$ are the signal fractions for voxel i .

As in Orton et al. [8], we sample from this with a Metropolis-Hastings algorithm. Proposed parameters are first sampled from Gaussian distributions as

$$x_{i,j}^* \sim N(x_{i,j}, w_{x_{i,j}}) \quad (13)$$

where $x_{i,j}$ is the current value of the parameter, $x_{i,j}^*$ is the proposed parameter value and $w_{x_{i,j}}$ is the variance of the proposal distribution, which should reflect the scale of the parameter and can be tuned for optimal algorithm performance.

The acceptance probability utilises the ratio of the posterior distributions for $x_{i,j}$ and $x_{i,j}^*$

$$\alpha(x_{i,j} \rightarrow x_{i,j}^*) = \min \left\{ 1, \frac{p(x_{i,j}^*|x_{i,-j}, \{f_k\}_{k=1}^{N_{comp}-1}, \theta_\mu, \Sigma)}{p(x_{i,j}|x_{i,-j}, \{f_k\}_{k=1}^{N_{comp}-1}, \theta_\mu, \Sigma)} \right\} \quad (14)$$

where the values on the right of the posterior are the current parameter values in the MCMC algorithm.

The signal fraction parameter MCMC moves are the same, except that the posterior distributions now contain the terms enforcing $\sum_{k=1}^{N_{comp}} f_k = 1$, i.e.

$$p(f_{i,k}|x_{i,1}, \dots, x_{i,J}, f_{i,-k}, \theta_\mu, \Sigma) \propto \begin{cases} [\mathbf{y}_i^T \mathbf{y}_i - (\mathbf{y}_i^T \mathbf{g}_i)^2 / \mathbf{g}_i^T \mathbf{g}_i]^{-N/2} & \text{if } \sum_{k=1}^{N_{comp}-1} f_k \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

where $f_{i,-k} = \{f_{i,1}, \dots, f_{i,k-1}, f_{i,k+1}, f_{i,K}\}$ are the other signal fractions apart from $f_{i,k}$. Again we sample proposed values as

$$f_{i,k}^* \sim N(f_{i,k}, w_{f_{i,k}}) \quad (16)$$

where $f_{i,k}$ is the current signal fraction. The acceptance probabilities are

$$\alpha(f_{i,k} \rightarrow f_{i,k}^*) = \min \left\{ 1, \frac{p(f_{i,k}^*|x_{i,1}, \dots, x_{i,J}, f_{i,-k}, \theta_\mu, \Sigma)}{p(f_{i,k}|x_{i,1}, \dots, x_{i,J}, f_{i,-k}, \theta_\mu, \Sigma)} \right\} \quad (17)$$

Metropolis-Hastings Acceptance Ratio. We tuned the Metropolis-Hastings jumping variances w_{θ_i} during the burn-in period to achieve an acceptance ratio that samples the posterior distribution efficiently. Following Orton et al. [8], at every 100 MCMC steps we applied the update rule

$$w_{\theta_i} = w_{\theta_i} 101 / (2(101 - R_{\theta_i})) \quad (18)$$

where R_{θ_i} is the number of times the proposed parameter update was accepted in the previous 100 steps. This scheme aims to adjust the jumping variances such that an acceptance rate of approximately 50% is achieved.

2.6 Models

The MCMC algorithm was tested using the ball-stick model, defined as

$$g(\theta, t) = f_{par} \exp(-bD_{par}(\mathbf{n} \cdot \mathbf{g})) + (1 - f_{par}) \exp(-bD_{iso}) \quad (19)$$

where b is the b-value, \mathbf{g} is the gradient direction and \mathbf{n} is the stick orientation, which is parameterised by two angles ϕ_1 and ϕ_2 constrained such that $\phi_1 \in (0, \pi)$ and $\phi_2 \in (-\pi, \pi)$. The signal fractions were constrained as $f_k \in (0.01 - 0.99)$ and the diffusivities as $D_{par}, D_{iso} \in (0.1 - 3) \mu\text{m}^2/\text{ms}$.

2.7 Algorithm implementation

Note that while all distributions have been presented in the linear scale, they were calculated in log-scale for numerical convenience. Parameter values were initialised with a voxelwise least squares fit, estimated using the Dmipy brute2fine option [1]. The MCMC algorithm was then run for 2000 steps with a burn-in of 1000 steps; weights were updated every 100 steps during the first half of the burn-in period. In our experience this was sufficient to sample the posterior distributions, and aligns with the work of Harms et al. [3]. We calculated model parameter posterior distributions and representative statistics from the 1000 MCMC samples after the burn-in. Parameter maps were generated using the mean of the posterior distributions in each voxel.

2.8 Data

To test the MCMC algorithm’s ability to infer correct model parameter values, we ran simulations using the Shepp-Logan phantom. We generated synthetic images with a matrix size of 128×128 and defined ground truth parameters in each major region (see Figure 3A, top row). We then simulated the signal in each voxel using Dmipy’s simulate_signal function with the same acquisition parameters as the Human Connectome Project (HCP) data (see below), added Gaussian noise to give a signal-to-noise ratio (SNR) of 10 in the $b = 0$ data, and ran the MCMC algorithm on these synthetic datasets. We perturbed initial parameter values to verify that the algorithm could find the global minimum. Bayesian priors were defined over the whole phantom excluding the background (i.e. one ROI).

We then applied our Bayesian model fitting approach on publicly-available data provided by the HCP WU-Minn Consortium (48 Subjects Test Retest Data Release, release date: Mar 01, 2017, available online at humanconnectome.org). Data from a single subject was used. The white matter (WM), cortical gray matter (GM), sub-cortical GM and ventricle ROIs derived from the Freesurfer

Algorithm 1 MCMC algorithm for Bayesian model fitting of a general microstructural model. This pseudocode describes the algorithm for a single ROI. For multiple ROIs, the algorithm is simply run separately on each ROI.

```

for voxels  $i = 1$  to  $i = N$  do
  Calculate initial values for voxelwise parameters:  $\{x_{i,j}\}_{j=1}^J$ ,  $\{f_{i,k}\}_{k=1}^{N_{comp}}$  with
  least squares estimation
end for
 $S \leftarrow$  number of MCMC steps
for MCMC steps  $s = 1$  to  $s = S$  do
   $\theta_{\mu}^{(s)} \leftarrow$  sample from Equation (10)
   $\Sigma^{(s)} \leftarrow$  sample from equation (11)
  for voxels  $i = 1$  to  $i = N$  do
    for non-signal fraction parameters  $j = 1$  to  $j = J$  do
       $\epsilon_{x_{i,j}} \leftarrow N(x_{i,j}^{(s-1)}, w_{x_{i,j}})$ 
       $x_{i,j}^* \leftarrow x_{i,j}^{(s-1)} + \epsilon_{x_{i,j}}$ 
      calculate  $\alpha(x_{i,j} \rightarrow x_{i,j}^*)$  from Equation (14)
       $u \leftarrow$  sample from  $\text{unif}(0, 1)$ 
      if  $u < \alpha$  then
         $x_{i,j}^{(s)} \leftarrow x_{i,j}^*$ 
      else
         $x_{i,j}^{(s)} \leftarrow x_{i,j}^{(s-1)}$ 
      end if
    end for
    for signal fraction parameters  $k = 1$  to  $k = N_{comp} - 1$  do
       $\epsilon_{f_{i,k}} \leftarrow N(f_{i,k}^{(s-1)}, w_{f_{i,k}})$ 
       $f_{i,k}^* \leftarrow f_{i,k}^{(s-1)} + \epsilon_{f_{i,k}}$ 
      calculate  $\alpha(f_{i,k} \rightarrow f_{i,k}^*)$  from Equation (17)
       $u \leftarrow$  sample from  $\text{unif}(0, 1)$ 
      if  $u < \alpha$  then
         $f_{i,k}^{(s)} \leftarrow f_{i,k}^*$ 
      else
         $f_{i,k}^{(s)} \leftarrow f_{i,k}^{(s-1)}$ 
      end if
    end for
     $f_{i,N_{comp}}^{(s)} \leftarrow 1 - \sum_{k=1}^{N_{comp}-1} f_{i,k}$ 
    if  $s \bmod 100 = 0$  then
      for voxels  $i = 1$  to  $i = N$  do
        Update  $\{w_{x_{i,j}}\}_{j=1}^J$  and  $\{w_{f_{i,k}}\}_{k=1}^{N_{comp}-1}$  using Equation (18)
      end for
    end if
  end for
end for

```

T_1 segmentations (these provided the best contrast between tissues) were transformed into diffusion space via linear and non-linear registration between the subject's T_1 -weighted image and the $b = 0$ dMRI data. The MCMC algorithm was applied with the Bayesian priors defined over these four ROIs.

3 Results

Figure 2 displays the output of several runs of the MCMC algorithm for a single voxel in the Shepp-Logan data. The MCMC chains and posterior distributions demonstrate that voxelwise parameter estimates converged to the ground truth value under a range of perturbations. Figure 3 compares the least squares and Bayesian parameter maps with the ground truth. The Bayesian approach more accurately replicated the ground truth and provided lower errors than least squares approach, particularly in low SNR cases.

Figure 4 shows the MCMC algorithm results on the HCP data. The Bayesian fit clearly removed some apparent outlier voxels when compared to the least squares fit (see arrows).

4 Discussion and Conclusions

In this work we present an extension to previous approaches that enables Bayesian hierarchical model fitting for a general microstructural model with arbitrary parameter constraints and regional priors. The algorithm is implemented by utilising and adapting the Dmipy software package, and is made publicly available

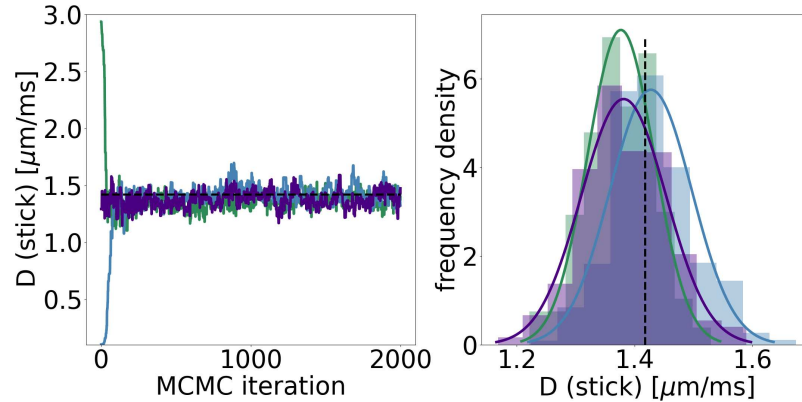


Fig. 2. MCMC output from the Bayesian ball-stick model fit on the Shepp-Logan phantom data. The left panel shows three MCMC chains for the stick parallel diffusivity in a single voxel; the initial parameter value in each run was given a different perturbation. The right panel shows the posterior distribution from each run, calculated on all samples after the burn-in of 1000 steps. The ground truth parameter value is indicated by the black lines.

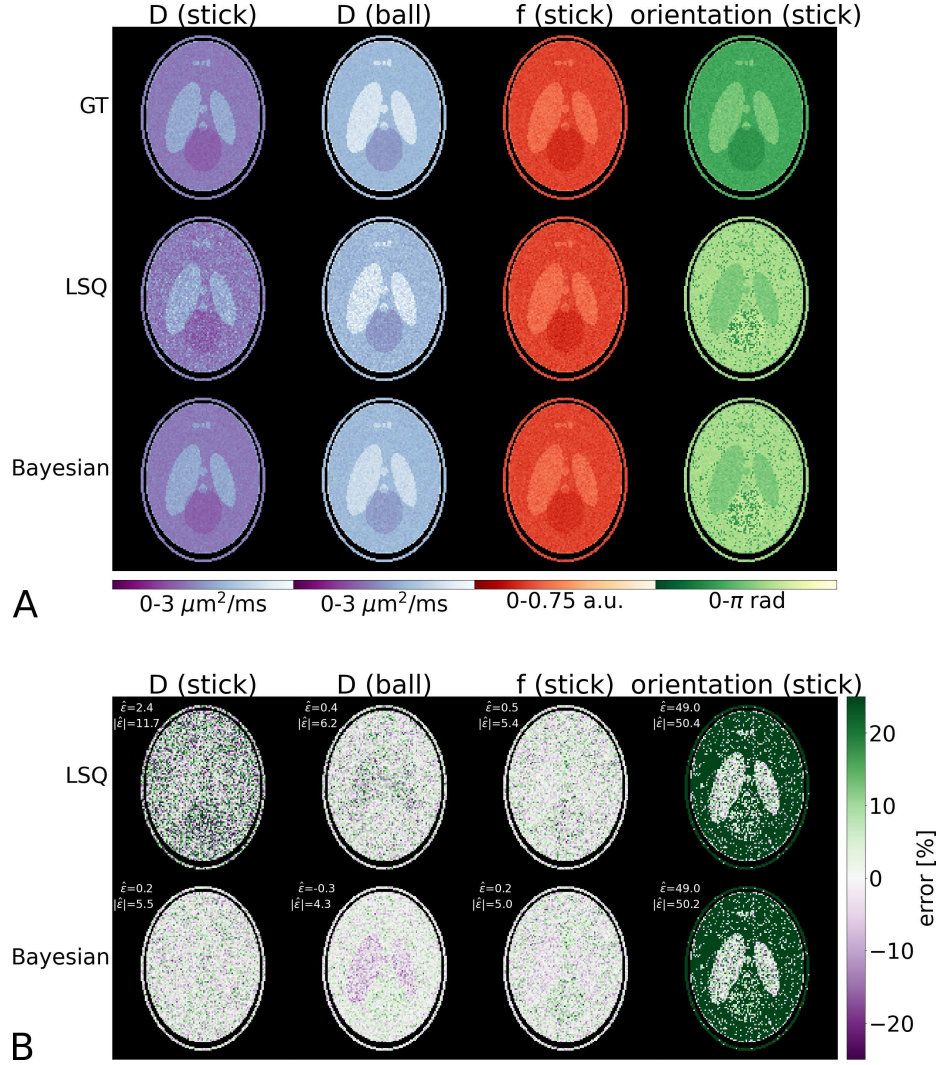


Fig. 3. Ball-stick model parameter maps in the Shepp-Logan phantom synthetic images. A. Ground truth (top row), least-squares derived (middle row) and Bayesian derived (bottom row) parameter maps for the stick parallel diffusivity (D_{par}) in $\mu\text{m}^2\text{s}^{-1}$, ball isotropic diffusivity (D_{iso}) in $\mu\text{m}^2\text{s}^{-1}$, stick signal fraction (f_{par}) and elevation orientation parameter (ϕ_1) in radians. B. Relative error maps for the least-squares fits (top row) and Bayesian fits (bottom row). Bayesian priors were defined over the whole image. The mean relative error ($\hat{\epsilon}$) and mean absolute relative error ($|\hat{\epsilon}|$) are displayed for each fitted parameter (both in %). Computation time for the Bayesian method was approximately 1 hour.

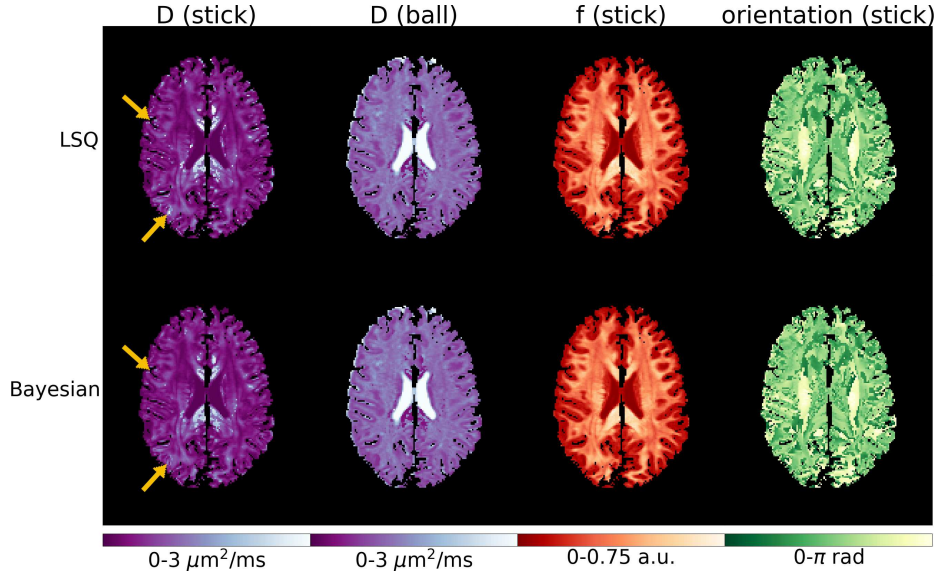


Fig. 4. Ball-stick parameter maps in the HCP data for the least squares fit (top row) and Bayesian fit (bottom row). Parameters as Figure 3. The Bayesian priors were defined over four ROIs as described in Methods section 2.8. Computation time for the Bayesian method was approximately 3.5 hours.

at github.com/PaddySlator/dmipy-bayesian. On synthetic data, we show that Bayesian fitting of the ball-stick model more accurately recovered ground truth maps than least squares fitting, particularly for parameters more susceptible to noise such as the stick parallel diffusivity. On HCP data, the algorithm reduced the appearance of apparent outlier voxels.

Although high SNR images from the HCP were used as the test data here, we anticipate the biggest gains of this approach will be seen in lower SNR data and more complex models. This may enable estimation of richer microstructural detail in a greatly reduced acquisition time.

The algorithm has limitations that motivate future work. The Bayesian approach assumes Gaussian noise, which may not be appropriate in all cases, particularly in low SNR cases. We also assumed that a Gaussian prior is suitable for all parameters; however, this may not be the optimal choice, particularly for orientation parameters where a flat prior may be more appropriate and could enable improved orientation parameter maps over LSQ (our current Bayesian implementation doesn't dramatically improve orientation maps, see Figures 3 and 4). Alternative prior choices are an avenue for future work. Utilising probabilistic segmentations, rather than the current hard-thresholded ROIs which may bias parameter estimation in partial volume voxels, is possible, but may complicate the MCMC inference. More complex model choices, as well as com-

parisons with alternative model fitting methods (e.g. Harms et al. [3]), should also be explored to better quantify the benefits of our Bayesian approach.

To conclude, we derive a general Bayesian hierarchical microstructural model and an MCMC algorithm for model inference given dMRI data. The algorithm, and corresponding open-source software, newly enables Bayesian model fitting for a wide range of microstructure imaging techniques.

5 Acknowledgements

This work was supported by EPSRC grants EP/M020533/1, EP/V034537/1 and EP/S031510/1, as well as the NIHR UCLH Biomedical Research Centre. The authors would like to thank the organisers of the Brainhack Micro2Macro 2021 hackathon, and the International Global Brainhack 2020.

References

1. Fick, R.H., Wassermann, D., Deriche, R.: The Dmipy Toolbox: Diffusion MRI Multi-Compartment Modeling and Microstructure Recovery Made Easy. *Frontiers in Neuroinformatics* **13**(October), 1–26 (2019)
2. Flouri, D., Owen, D., Aughwane, R., Mufti, N., Maksym, K., Sokolska, M., Kendall, G., Bainbridge, A., Atkinson, D., Vercauteren, T., Ourselin, S., David, A.L., Melbourne, A.: Improved fetal blood oxygenation and placental estimated measurements of diffusion-weighted MRI using data-driven Bayesian modeling. *Magnetic Resonance in Medicine* **83**(6), 2160–2172 (2020)
3. Harms, R.L., Roebroek, A.: Robust and Fast Markov Chain Monte Carlo Sampling of Diffusion MRI Microstructure Models. *Frontiers in Neuroinformatics* p. 97 (dec)
4. Kaden, E., Kelm, N.D., Carson, R.P., Does, M.D., Alexander, D.C.: Multi-compartment microscopic diffusion imaging. *NeuroImage* **139**, 346–359 (2016)
5. Koh, D.M., Collins, D.J., Orton, M.R.: Intravoxel Incoherent Motion in Body Diffusion-Weighted MRI: Reality and Challenges. *American Journal of Roentgenology* (6), 1351–1361 (jun)
6. Le Bihan, D., Breton, E., Lallemand, D., Aubin, M.L.L., Vignaud, J., Laval-Jeantet, M.: Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* (2), 497–505 (aug)
7. Novikov, D.S., Fieremans, E., Jespersen, S.N., Kiselev, V.G.: Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR in Biomedicine* (4), e3998 (apr)
8. Orton, M.R., Collins, D.J., Koh, D.M.M., Leach, M.O.: Improved intravoxel incoherent motion analysis of diffusion weighted imaging by data driven Bayesian modeling. *Magnetic Resonance in Medicine* (1), 411–420
9. Panagiotaki, E., Chan, R.W., Dikaios, N., Ahmed, H.U., O’Callaghan, J., Freeman, A., Atkinson, D., Punwani, S., Hawkes, D.J., Alexander, D.C.: Microstructural characterization of normal and malignant human prostate tissue with vascular, extracellular, and restricted diffusion for cytometry in tumours magnetic resonance imaging. *Investigative radiology* **50**(4), 218–227 (2015)
10. Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C.: NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* (4), 1000–1016