



HAL
open science

Inverting brain grey matter models with likelihood-free inference: a tool for trustable cytoarchitecture measurements

Maëliss Jallais, Pedro L C Rodrigues, Alexandre Gramfort, Demian Wassermann

► To cite this version:

Maëliss Jallais, Pedro L C Rodrigues, Alexandre Gramfort, Demian Wassermann. Inverting brain grey matter models with likelihood-free inference: a tool for trustable cytoarchitecture measurements. 2021. hal-03417406v1

HAL Id: hal-03417406

<https://inria.hal.science/hal-03417406v1>

Preprint submitted on 8 Nov 2021 (v1), last revised 2 May 2022 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Inverting brain grey matter models with likelihood-free inference: a tool for trustable cytoarchitecture measurements

Maëliss Jallais

Université Paris-Saclay, Inria, CEA, Palaiseau, France

maeliss.jallais@inria.fr

Pedro L. C. Rodrigues

Université Paris-Saclay, Inria, CEA, Palaiseau, France

pedro.rodrigues@inria.fr

Alexandre Gramfort

Université Paris-Saclay, Inria, CEA, Palaiseau, France

alexandre.gramfort@inria.fr

Demian Wassermann

Université Paris-Saclay, Inria, CEA, Palaiseau, France

demian.wassermann@inria.fr

Abstract

1 Effective characterisation of the brain grey matter cytoarchitecture with quantitative sensi-
 2 tivity to soma density and volume remains an unsolved challenge in diffusion MRI (dMRI).
 3 Solving the problem of relating the dMRI signal with cytoarchitectural characteristics calls
 4 for the definition of a mathematical model that describes brain tissue via a handful of
 5 physiologically-relevant parameters and an algorithm for inverting the model. To address
 6 this issue, we propose a new forward model, specifically a new system of equations, requir-
 7 ing a few relatively sparse b -shells. We then apply modern tools from Bayesian analysis
 8 known as likelihood-free inference (LFI) to invert our proposed model. As opposed to other
 9 approaches from the literature, our algorithm yields not only an estimation of the param-
 10 eter vector θ that best describes a given observed data point \mathbf{x}_0 , but also a full posterior
 11 distribution $p(\theta|\mathbf{x}_0)$ over the parameter space. This enables a richer description of the
 12 model inversion, providing indicators such as credible intervals for the estimated paramet-
 13 ers and a complete characterization of the parameter regions where the model may present
 14 indeterminacies. We approximate the posterior distribution using deep neural density esti-
 15 mators, known as normalizing flows, and fit them using a set of repeated simulations from
 16 the forward model. We validate our approach on simulations using `dmipy` and then apply
 17 the whole pipeline on two publicly available datasets.

18 **Keywords:** Diffusion MRI, Brain Microstructure, Likelihood-Free Inference.

19 1. Introduction

20 Obtaining quantitative measurements of brain grey matter microstructure with a dedicated
 21 soma representation is a growing field of interest in the diffusion MRI (dMRI) community
 22 (Palombo et al., 2020; Jelescu et al., 2021). Unlike histology, dMRI permits to quantify
 23 brain tissue characteristics non-invasively and could, for example, help understanding de-
 24 mentia and cognitive deficits (Douaud et al., 2013). However, current methods require
 25 demanding acquisitions with several q -shells (equivalently b -shells) and rely on non-linear
 26 models for which several parameter values may yield the same observation, also known as
 27 parameter indeterminacy (Jelescu et al., 2016; Novikov et al., 2018b). This leads to numer-
 28 ically unstable solutions which are hard to interpret. Another major challenge is the fact

29 that quantifying brain tissue microstructure directly from the dMRI signal is an inherently
 30 difficult task, because of the large dimensionality of the collected data and its stochastic
 31 nature (Fick et al., 2016). For instance, each one of the $\approx 10^4$ voxels in a dMRI acquisition
 32 is sampled along nine dimensions, three directions in the q -space for each direction in the
 33 \mathbb{R}^3 spatial domain, which leads to very large data structures to process.

34 Many works have tackled the non-linear inverse problem of inferring brain tissue pa-
 35 rameters from dMRI signal. A popular solution is NODDI (e.g. Zhang et al., 2012), which
 36 stabilises the solution by imposing constraints on model parameters that are not biologi-
 37 cally plausible (Novikov et al., 2018b). However, this has the effect of biasing the parameter
 38 estimation and the inverse problem remains largely degenerate (Jelescu et al., 2016). There
 39 has also been some interest in applying methods from the machine learning literature to
 40 solve the inverse problem. This is the case of the SANDI model (Palombo et al., 2020), in
 41 which the authors employ random forest regressors. Although the method provides rather
 42 acceptable accuracy in real-case scenarios, it can only output one set of tissue parameters
 43 for a given observed dMRI signal, masking, therefore, other biologically plausible solutions
 44 that could generate the same observed signal.

45 To overcome such limitations, we present three contributions. First, we use a three-
 46 compartment model for brain tissue and introduce a new parameter that jointly encodes
 47 soma radius and inner diffusivity without imposing constraints on these values. This new
 48 parameter reduces indeterminacies in the model and has relevant physiological interpre-
 49 tations. Second, we present a method to fit the model through summary features of the
 50 dMRI signal based on a large and small q -value analysis using boundary approximations.
 51 These rotationally-invariant features relate directly to the tissue parameters and enable us
 52 to invert the model without manipulating the raw dMRI signals. Such summary statis-
 53 tics ensure a stable solution of the parameter estimations, as opposed to the indeterminate
 54 models used in Zhang et al. (2012); Palombo et al. (2020). Third, we employ modern tools
 55 from Bayesian analysis known as likelihood-free inference (LFI, (Cranmer et al., 2020))
 56 to solve our non-linear inverse problem under a probabilistic framework and determine the
 57 posterior distribution of the fitted parameters. Such approach offers a full description of the
 58 solution landscape and can point out degeneracies, as opposed to the usual deterministic
 59 least-squares based solution (Jelescu et al., 2016; Novikov et al., 2018b).

60 2. Related Works

61 Current brain tissue models are predominantly based on the two compartment Standard
 62 Model (SM) (Zhang et al., 2012; Novikov et al., 2018a). Recent evidence shows that the SM,
 63 mainly used in white matter, does not hold for grey matter microstructure analysis (Veraart
 64 et al., 2020). Several assumptions aim at explaining this issue such as increased permeability
 65 in neurite membranes (Veraart et al., 2020), or curvy projections along with longer pulse
 66 duration (Novikov et al., 2018a). We follow the hypothesis that the SM doesn't hold due to
 67 an abundance of cell bodies in grey matter (Palombo et al., 2020). Our proposed biophysical
 68 model is then based on three compartments (Palombo et al., 2020): neurites, somas, and
 69 extra-cellular space (ECS). Despite its increased complexity, the main advantage of such
 70 model is the possibility to jointly estimate the characteristic features of each compartment.

71 Inferring parameters of the brain tissue model directly from dMRI signals has proven to
 72 be a very challenging task and has motivated the development of new approaches to reduce
 73 the dimensionality of the data to be processed. For instance, Novikov et al. (2018b) proposes
 74 the LEMONADE system of equations, based on a Taylor expansion of the diffusion signal
 75 and a set of rotationally invariant moments. In this work, we extend the LEMONADE
 76 equations to a setting with a three compartment model and further develop the method so
 77 to extract more features from the observed signal. We call the resulting vector of features
 78 defined by these quantities the ‘summary statistics’ of the dMRI signal.

79 The usual way of applying a Bayesian approach to solve non-linear inverse problems (Stu-
 80 art, 2010) is to define two quantities: a prior distribution encoding initial knowledge of
 81 the parameter values (e.g. intervals which are physiologically relevant) and the likelihood
 82 function of the forward model being studied. One can then either obtain an analytic ex-
 83 pression of the posterior distribution via Bayes’ formula or use a Markov-Chain Monte
 84 Carlo (MCMC) procedure to numerically sample the posterior distribution (Gelman et al.,
 85 2013). However, the likelihood function of complex models such as the one that we consider
 86 here is often very hard to obtain and makes the Bayesian approach rather challenging to
 87 use. Likelihood-free inference (LFI) bypasses this bottleneck by recurring to several simu-
 88 lations of the forward model using different parameters and learning an approximation to
 89 the posterior distribution from these examples (Cranmer et al., 2020).

90 The first contributions on LFI are also known as approximate Bayesian computation
 91 (ABC) and have been applied to invert models from ecology, population genetics, and
 92 epidemiology (Sisson, 2018). Some of the limitations of these techniques include the large
 93 number of simulations required for the posterior estimations and the need of defining a
 94 distance function to compare the results of two simulations. Recently, there has been a
 95 growing interest in the machine learning community in improving the limitations of ABC
 96 methods through deep generative modeling, i.e. neural network architectures specially
 97 tailored to approximate probability density functions from a set of examples (Goodfellow
 98 et al., 2016). Normalizing flows (Papamakarios et al., 2019) are a particular class of such
 99 neural networks that have demonstrated promising results for likelihood-free inference in
 100 different research fields (Cranmer et al., 2020; Gonçalves et al., 2020; Greenberg et al.,
 101 2019).

102 3. Theory

103 In this section, we present the theoretical background underlying our three main contribu-
 104 tions. **Section 3.1** describes the three-compartment model for brain grey matter. We also
 105 introduce a new parameter that captures both soma radius and inner diffusivity and avoids
 106 the usual indeterminacy in estimating them separately. **Section 3.2** presents the summary
 107 statistics used to reduce the dimensionality of the dMRI signal into a 7-dimensional feature
 108 vector that can be used to determine the physiological parameters of interest. **Section 3.3**
 109 describes the Bayesian approach used to solve our non-linear inverse problem and obtain
 110 a posterior distribution for the physiological parameters that best describe a given dMRI
 111 signal.

112 3.1 Modeling the Brain grey Matter with a 3-compartment Model

113 To characterize cortical cytoarchitecture, we propose a method that relates the diffusion
 114 MRI signal to specific tissue parameters. To that aim, we first define a forward model
 115 based on a biophysical modeling of brain grey matter.

116 Research in histology has demonstrated that grey matter is composed of neurons em-
 117 bedded in a fluid environment. Each neuron is composed of a soma, corresponding to the
 118 cellular body, surrounded by neurites connecting neurons together. Following this tissue
 119 biophysical composition, we model the grey matter tissue as three-compartmental (Palombo
 120 et al., 2020), moving away from the usual standard model (SM) designed for white matter.
 121 We further assume that our acquisition protocol is insensitive to exchanges between the
 122 compartments, i.e. molecules moving from one compartment to another have a negligible
 123 influence on the signal (Palombo et al., 2020). Many works include also a dot compartment
 124 into the SM with zero apparent diffusivity and no exchange. However, we have not consid-
 125 ered such assumption, because its presence has been considered very unlikely in grey matter
 126 by previous works (Veraart et al., 2019; Tax et al., 2020). The observed diffusion signal is
 127 considered as a convex mixture of signals arising from somas, neurites, and extra-cellular
 128 space (ECS). Unlike white matter-centric methods (Jelescu and Budde, 2017), we are not
 129 interested in the fiber orientation and only estimate orientation-independent parameters.
 130 This enables us to work on the direction-averaged dMRI signal, denoted $\bar{S}(q)$, known as the
 131 power averaged signal. This consideration mainly matters for neurites, as their signal is not
 132 isotropic, as opposed to the proposed model for somas and ECS. Our direction-averaged
 133 grey matter signal model is then:

$$\frac{\bar{S}(q)}{S(0)} = f_n \bar{S}_{\text{neurites}}(q, D_n) + f_s \bar{S}_{\text{somas}}(q, D_s, r_s) + f_{\text{ECS}} \bar{S}_{\text{ECS}}(q, D_e) . \quad (1)$$

134 In this equation, f_n , f_s , and f_{ECS} represent signal fractions for neurites, somas, and ECS
 135 respectively ($f_n + f_s + f_{\text{ECS}} = 1$). Note that the relative signal fractions do not correspond
 136 to the relative volume fractions of the tissue compartments as they are also modulated
 137 by different T2 values (Novikov et al., 2018a). D_n corresponds to axial diffusivity inside
 138 neurites, while D_s and D_e correspond to somas and extra-cellular diffusivities. r_s is the
 139 average soma radius within a voxel. This model is the same as the model SANDI proposed
 140 by Palombo et al. (2020), with $f_n = f_{icfin}$, $f_s = f_{icfis}$ and $f_{\text{ECS}} = f_{ec}$. We use q -values
 141 for more readability and harmonization throughout the paper, but a direct conversion to
 142 b -values is also possible, using $b = (2\pi q)^2 \tau$ with $\tau = \Delta - \delta/3$.

143 We now review the model for each compartment, to make explicit the impact of each
 144 parameter on the diffusion MRI signal.

145 **Neurite compartment.** Neurites, as in the SM, are modeled as 0-radius imperme-
 146 able cylinders (“sticks”), with effective diffusion along the parallel axis, and a negligible
 147 radial intra-neurite diffusivity. In our acquisition setting, this model has been shown to
 148 hold (Veraart et al., 2020). Its direction averaged signal is (Veraart et al., 2020):

$$\bar{S}_{\text{neurites}}(q) \simeq \frac{1}{4\sqrt{\pi\tau D_n}} \cdot q^{-1} . \quad (2)$$

149 **Soma compartment.** Somas are modeled as spheres, whose signal can be computed
 150 using the GPD approximation (Balinov et al., 1993):

$$-\log \bar{S}_{\text{somas}}(q) = C(r_s, D_s) \cdot q^2 . \quad (3)$$

We exploit this relation here to extract a parameter $C_s = C(r_s, D_s)[m^2]$ which, at fixed diffusivity D_s , is modulated by the radius of the soma r_s :

$$C(r_s, D_s) = \frac{2}{D_s \delta^2} \sum_{m=1}^{\infty} \frac{\alpha_m^{-4}}{\alpha_m^2 r_s^2 - 2} \cdot \left(2\delta - \frac{2 + e^{-\alpha_m^2 D_s (\Delta - \delta)} - e^{-\alpha_m^2 D_s \delta} - e^{-\alpha_m^2 D_s \Delta} + e^{-\alpha_m^2 D_s (\Delta + \delta)}}{\alpha_m^2 D_s} \right) ,$$

151 where α_m is the m th root of $(\alpha r_s)^{-1} J_{\frac{3}{2}}(\alpha r_s) = J_{\frac{5}{2}}(\alpha r_s)$, with $J_n(x)$ the Bessel functions of
 152 the first kind. In certain specific cases, C_s has a simpler and more interpretable expression.
 153 For instance, when we consider a narrow pulse regime, with small τ or large r_s , C_s loses its
 154 dependence on r_s (Balinov et al., 1993) and we obtain:

$$C_s = D_s \tau . \quad (4)$$

155 In the Neuman (wide pulse) regime, i.e. when $D_s \Delta \gg r_s^2$ and $D_s \delta \ll 1$, C_s becomes only
 156 dependent on the soma radius (Murday and Cotts, 1968):

$$C_s = \frac{1}{5} r_s^2 . \quad (5)$$

157 These two approximations permits us to better interpret the parameter C_s .

158 **Extra-cellular space compartment.** The extra-cellular space is approximated as
 159 isotropic Gaussian diffusion, i.e. a mono-exponential diffusion signal with a scalar diffusion
 160 constant D_e , which reflects the molecular viscosity of the fluid. This approximation assumes
 161 that the ECS is fully connected. The approximation is therefore:

$$-\log(\bar{S}_{\text{ECS}}(q)) = (2\pi q)^2 \tau D_e . \quad (6)$$

162 Because of the geometry of the problem, we estimate D_e as equal to one third of the diffu-
 163 sivity in the ventricles (considered as free diffusivity), given the same metabolic composition
 164 of the extracellular fluid and ventricles (Vincent et al., 2021).

165 3.2 An Invertible 3-compartment Model: dMRI Summary Statistics

166 The tissue model presented in Section 3.1 enables us to relate the dMRI signal with param-
 167 eters representing grey matter tissue microstructure. However, solving the inverse problem
 168 directly from Eq. (1) is a difficult task, leading to indeterminacies and poorly estimated
 169 parameters with large variability. Current methods addressing this issue have not studied
 170 its stability (Palombo et al., 2020) but simpler models with only two compartments have
 171 been shown to be indeterminate (Novikov et al., 2018a).

172 To produce a method which addresses this indeterminacy, we introduce rotationally
 173 invariant summary statistics to describe the dMRI signal. The goal being to reduce the di-
 174 dimensionality of the data at hand and representing all the relevant information for statistical

175 inference with a few features. We then solve the inverse problem efficiently using Bayesian
 176 inference as described in Section 3.3. These dMRI-based summary statistics are extracted
 177 from our proposed model presented in Section 3.1 via the following analysis of the dMRI
 178 signal on the boundaries of large and small q -value cases.

179 **Large q -value approximation: RTOP.** We compute a q -bounded return-to-the-
 180 origin probability (RTOP), which measures the restrictions of the diffusing fluid molecule
 181 motion and gives us information about the structure of the media (Mitra et al., 1995):

$$\text{RTOP}(q) = 4\pi \int_0^q \frac{\bar{S}(\eta)}{S(0)} \eta^2 d\eta . \quad (7)$$

182 For q large enough, the RTOP in our 3-compartment model in Eq. (1) yields a soma and
 183 extra-cellular signal which converges towards a constant value in q , while the neurites'
 184 contribution becomes quadratic in q . In this case, RTOP becomes:

$$\text{RTOP}(q) = \underbrace{f_s \left(\frac{\pi}{C_s} \right)^{3/2}}_{a_{\text{fit}}} + \frac{f_{\text{ECS}}}{8(\pi\tau D_e)^{3/2}} + \underbrace{\frac{f_n}{2} \cdot \sqrt{\frac{\pi}{\tau D_n}}}_{b_{\text{fit}}} \cdot q^2 + \gamma q^3 , \quad (8)$$

185 in which the last term of the equation, γq^3 , is a nuisance parameter that describes a constant
 186 noise in the direction averaged signal. By accurately estimating the second derivative of
 187 RTOP(q) at q large enough, we can solve the coefficients of interest of the polynomial in
 188 Eq. (8): a_{fit} and b_{fit} . We do this efficiently by casting it as an overdetermined ordinary least
 189 squares problem which has a unique solution.

190 **Small q -value approximation: Spiked LEMONADE.** We propose a second ap-
 191 proximation, based on a moment decomposition for small q -values (Novikov et al., 2018b):

$$\frac{S_{\hat{\mathbf{g}}}(q)}{S(0)} = 1 - b(q) M_{i_1 i_2}^{(2)} g_1 g_2 + \frac{b(q)^2}{2!} M_{i_1 \dots i_4}^{(4)} g_1 \dots g_4 - \dots , \quad b(q) = (2\pi q)^2 \tau , \quad (9)$$

193 where i_k are the directional basis of the tensors M , $g_k = i_k \cdot \hat{\mathbf{g}} \in \mathbb{R}^3$, and $\hat{\mathbf{g}}$ the unit direc-
 194 tion of the dMRI acquisition. From the moment tensors of this decomposition, LEMON-
 195 ADE (Novikov et al., 2018b) extracts rotational invariant scalar indices $M^{(i),j}$, $i, j \in \{0, 2, 4, \dots\}$.
 196 These quantify white matter microstructure by plugging the 2-compartment SM into Eq. (9) (see
 197 Novikov et al., 2018b, app. C).

198 In this work, we extended LEMONADE to our 3-compartment model presented in
 199 Section 3.1 by plugging Eq. (1) into Eq. (9) and performing tedious arithmetic. This results
 200 in the following equation system, which now includes the soma parameter C_s , relating the
 201 dMRI signal to grey matter microstructure:

$$\begin{cases} M^{(2),0} = f_n D_n + 3f_s \frac{C_s}{(2\pi)^2 \tau} + 3f_{\text{ECS}} D_e \\ M^{(2),2} = f_n D_n p_2 \\ M^{(4),0} = f_n D_n^2 + 5f_s \left(\frac{C_s}{(2\pi)^2 \tau} \right)^2 + 5f_{\text{ECS}} D_e^2 \\ M^{(4),2} = f_n D_n^2 p_2 \end{cases} \quad (10)$$

202 where p_2 is a scalar measure of neurite orientation dispersion (Novikov et al., 2018b).

203 Note that only the shells with $b(q) \leq 2.5 \text{ ms } \mu\text{m}^{-2}$ are used, to get an unbiased estimation
 204 of the rotational invariant moments $M^{(2),0}$, $M^{(2),2}$, $M^{(4),0}$ and $M^{(4),2}$.

205 **Complete system.** Combining equations (8) and (10) and adding the constraint that
 206 the fractions for the three compartments sum to one, we obtain a non-linear system of 7
 207 equations and 7 unknowns. Following Menon et al. (2020), we assume D_e nearly-constant
 208 per subject acquisition and estimated it as one third of the mean diffusivity in the subject's
 209 ventricles (Vincent et al., 2021). This assumption allows us to drop an unknown from the
 210 system, use D_e as a reference diffusivity and render our variables unitless as $D_n^u = \frac{D_n}{D_e}$ and
 211 $C_s^u = \frac{C_s}{(2\pi)^2 \tau D_e}$, obtaining:

Spiked LEMONADE	RTOP
Small q -values	Large q -values
$\begin{cases} \frac{M^{(2),0}}{D_e^e} = f_n D_n^u + 3f_s C_s^u + 3f_{\text{ECS}} \\ \frac{M^{(2),2}}{D_e^e} = f_n D_n^u \cdot p_2 \\ \frac{M^{(4),0}}{D_e^2} = f_n D_n^{u2} + 5f_s C_s^{u2} + 5f_{\text{ECS}} \\ \frac{M^{(4),2}}{D_e^2} = f_n D_n^{u2} \cdot p_2 \end{cases}$	$\begin{cases} a_{\text{fit}} (\tau D_e)^{3/2} = \frac{f_s}{8(\pi C_s^u)^{3/2}} \\ + \frac{f_{\text{ECS}}}{8\pi^{3/2}} \\ b_{\text{fit}} (\tau D_e)^{1/2} = \frac{f_n}{2} \sqrt{\frac{\pi}{D_n^u}} \end{cases}$
$f_n + f_s + f_{\text{ECS}} = 1$	

213 3.3 Inverting the model with Bayesian inference

Our main goal is to determine the values of the parameter vector

$$\boldsymbol{\theta} = (D_n^u, C_s^u, p_2, f_s, f_n, f_{\text{ECS}}) \in \mathbb{R}^6$$

that best explain a given observed dMRI signal d . Because of the high-dimensionality of this kind of signal and the difficulties in obtaining stable estimates of $\boldsymbol{\theta}$ directly from it, we recur to the set of summary features defined in Section 3.2,

$$\boldsymbol{x} = \left(\frac{M^{(2),0}}{D_e}, \frac{M^{(2),2}}{D_e}, \frac{M^{(4),0}}{D_e^2}, \frac{M^{(4),2}}{D_e^2}, a_{\text{fit}}(\tau D_e)^{3/2}, b_{\text{fit}} \sqrt{\tau D_e} \right) \in \mathbb{R}^7,$$

214 and make the assumption that it carries all the information necessary for determining the $\boldsymbol{\theta}_0$
 215 having generated a given dMRI signal \boldsymbol{S}_0 . We denote the relation between these quantities
 216 as

$$\boldsymbol{x} = \mathcal{M}(\boldsymbol{\theta}) + \boldsymbol{n}, \quad (11)$$

217 where $\mathcal{M} : \mathbb{R}^6 \rightarrow \mathbb{R}^7$ is a multivariate function that implements the system of equations
 218 described in Section 3.2 and \boldsymbol{n} is an additive noise capturing the imperfections of our
 219 modelling procedure, the limitations of the summary statistics, and the measurement noise.

220 **The Bayesian formalism.** We interpret the inverse problem of inferring the parame-
 221 ters that best describe a given observed summary feature vector \boldsymbol{x}_0 as that of determining
 222 the posterior distribution of $\boldsymbol{\theta}$ given an observation \boldsymbol{x}_0 . By first choosing a prior distribution
 223 $p(\boldsymbol{\theta})$ describing our initial knowledge of the parameter values, we may use Bayes' theorem
 224 to write

$$p(\boldsymbol{\theta}|\boldsymbol{x}_0) = \frac{p(\boldsymbol{x}_0|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{x}_0)}, \quad (12)$$

225 where $p(\mathbf{x}_0|\boldsymbol{\theta})$ is the likelihood of the observed data point and $p(\mathbf{x}_0)$ is a normalizing con-
 226 stant, commonly referred to as the evidence of the data. Note that such a probabilistic
 227 approach returns not only which $\boldsymbol{\theta}$ best fits the observed data (i.e. the parameter that
 228 maximizes the posterior distribution), but the full posterior distribution $p(\boldsymbol{\theta}|\mathbf{x}_0)$. The lat-
 229 ter can be possibly multimodal or flat, which would indicate the difficulty of summarizing
 230 it with a unique maximum.

231 **Bypassing the likelihood function.** Despite its apparent simplicity, it is usually
 232 difficult to use Eq. (12) to determine the posterior distribution, since the likelihood function
 233 for data points generated by complex non-linear models is often hard to write. To avoid
 234 such difficulty, we directly approximate the posterior distribution using a conditional density
 235 estimator, i.e. a family of conditional p.d.f. approximators $q_\phi(\boldsymbol{\theta}|\mathbf{x})$ parametrized by ϕ and
 236 that takes $\boldsymbol{\theta}$ (the parameter) and \mathbf{x} (the observation) as input arguments. Our posterior
 237 approximation is then obtained by minimizing its average Kullback-Leibler divergence with
 238 respect to the conditional density estimator for different choices of \mathbf{x} , as per (Papamakarios
 239 and Murray, 2016)

$$\min_{\phi} \mathcal{L}(\phi) \quad \text{with} \quad \mathcal{L}(\phi) = \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})} [D_{\text{KL}}(p(\boldsymbol{\theta}|\mathbf{x}) \| q_\phi(\boldsymbol{\theta}|\mathbf{x}))] , \quad (13)$$

240 which can be rewritten as

$$\begin{aligned} \mathcal{L}(\phi) &= \int D_{\text{KL}}(p(\boldsymbol{\theta}|\mathbf{x}) \| q_\phi(\boldsymbol{\theta}|\mathbf{x})) p(\mathbf{x}) d\mathbf{x} , \\ &= - \iint \log (q_\phi(\boldsymbol{\theta}|\mathbf{x})) p(\boldsymbol{\theta}|\mathbf{x}) p(\mathbf{x}) d\boldsymbol{\theta} d\mathbf{x} + C , \\ &= - \iint \log (q_\phi(\boldsymbol{\theta}|\mathbf{x})) p(\mathbf{x}, \boldsymbol{\theta}) d\boldsymbol{\theta} d\mathbf{x} + C , \\ &= -\mathbb{E}_{(\mathbf{x}, \boldsymbol{\theta}) \sim p(\mathbf{x}, \boldsymbol{\theta})} [\log (q_\phi(\boldsymbol{\theta}|\mathbf{x}))] + C , \end{aligned} \quad (14)$$

241 where C is a constant that does not depend on ϕ . Note, however, that in practice we
 242 actually consider a N -sample Monte-Carlo approximation of the loss function given by

$$\mathcal{L}(\phi) \approx \mathcal{L}^N(\phi) = -\frac{1}{N} \sum_{i=1}^N \log \left(q_\phi(\boldsymbol{\theta}_i|\mathbf{x}_i) \right) , \quad (15)$$

243 where the N data points $(\boldsymbol{\theta}_i, \mathbf{x}_i)$ are sampled from the joint distribution with $\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta})$
 244 and $\mathbf{x}_i \sim p(\mathbf{x}|\boldsymbol{\theta}_i)$. We can then use stochastic gradient descent to obtain a set of param-
 245 eters ϕ which minimizes \mathcal{L}^N . If the class of conditional density estimators is sufficiently
 246 expressive, it is possible to show that the minimizer of Eq. (15) converges to $p(\boldsymbol{\theta}|\mathbf{y})$ when
 247 $N \rightarrow \infty$ (Greenberg et al., 2019). Note, also, that the parametrization ϕ that we obtain
 248 by the end of the optimization procedure yields a posterior which is amortized for different
 249 choices of \mathbf{x} . Thus, for a specific observation \mathbf{x}_0 we may simply write $q_\phi(\boldsymbol{\theta}|\mathbf{x}_0)$ to get an
 250 approximation of $p(\boldsymbol{\theta}|\mathbf{x}_0)$.

251 **Neural density estimators.** In this work, our conditional p.d.f. approximators belong
 252 to a class of neural networks called normalizing flows (Papamakarios et al., 2019). These
 253 flows are invertible functions capable of transforming vectors generated by a simple base
 254 distribution (e.g. the standard multivariate Gaussian distribution) into an approximation

255 of the true posterior distribution. An important advantage of normalizing flow versus
 256 other p.d.f. approximators such as generative adversarial network (GAN (Goodfellow et al.,
 257 2014)) and variational auto-encoders (VAE (Kingma and Welling, 2014)) is that it provides
 258 both the likelihood of any sample point of interest and it is also straightforward to sample
 259 new data points from it. Furthermore, certain classes of normalizing flows can be shown
 260 to be universal approximators of probability density functions. We refer the reader to
 261 Papamakarios et al. (2019) for more information on the different types of normalizing flows.

262 4. Materials and methods

263 This section presents the technical details on how we have implemented our theoretical con-
 264 tributions and describes the simulated and real datasets used in the numerical illustrations.

265 4.1 The likelihood-free inference setup

266 Using a likelihood-free inference approach for inverting the 3-compartment model relating
 267 tissue parameters θ and dMRI summary statistics \mathbf{x} relies on four important aspects:

- 268 (1) **The forward model.** As explained in Section 3.3, we obtain an approximation of the
 269 amortized posterior distribution using a dataset containing several paired examples
 270 of a parameter θ_i and its corresponding summary statistics \mathbf{x}_i . In what follows, we
 271 adopt the usual assumptions from the inverse problems literature and consider the
 272 additive noise \mathbf{n} from Eq. (11) small enough to be ignored in the data generation, so
 273 that we have $\mathbf{x}_i \approx \mathcal{M}(\theta_i)$.
- 274 (2) **Prior distribution.** The simplest way of defining a prior distribution $p(\theta)$ in our
 275 setting is to use an uniform distribution with limits within physiologically relevant
 276 intervals for each parameter. From Section 3.1, we know that the fractions f_s , f_n , and
 277 f_{ECS} have values between 0 and 1 and all sum up to one. To encode this information
 278 in $p(\theta)$, we define new parameters k_1 and k_2 and relate them with the fractions by

$$f_n = k_2 \sqrt{k_1}, \quad f_s = \sqrt{k_1}(1 - k_2), \quad f_{\text{ECS}} = 1 - \sqrt{k_1}. \quad (16)$$

279 In this way, whenever we want to generate a prior sample for $(f_n, f_s, f_{\text{ECS}})$ we first
 280 generate a sample of $k_1, k_2 \sim \mathcal{U}([0, 1])$ and then transform them according to (16) to
 281 get a set of fractions which is uniformly sampled in the region $\{f_n, f_s, f_{\text{ECS}} \in [0, 1] :$
 282 $f_n + f_s + f_{\text{ECS}} = 1\}$. We follow the usual assumption that the diffusivity of the com-
 283 partments are inferior or equal to the self-diffusion coefficient of free water, which is
 284 $3 \mu\text{m ms}^{-2}$ (Li et al., 2016). We fix, therefore, the interval for neurite diffusivity (D_n)
 285 as between 10^{-5} and $3 \mu\text{m ms}^{-2}$. The newly introduced parameter C_s is parametrized
 286 by soma radius and diffusivity. To account for a soma radius comprised between 2
 287 and $30 \mu\text{m}$ (Palombo et al.) and a diffusivity range as previously defined, we used
 288 the C_s interval $[50, 2500] \mu\text{m}^2$. Parameter p_2 , which measures the dispersion of neu-
 289 rites orientation, is comprised in the interval $[0, 1]$, with 1 indicating an anisotropic
 290 orientation distribution function.

- 291 (3) **Posterior approximator.** We use an autoregressive architecture for normalizing
 292 flows implemented via the masked autoencoder for distribution estimation (MADE) (Ger-

main et al., 2015). We follow the same setup from Greenberg et al. (2019) for LFI problems, stacking five MADEs, each with two hidden layers of 50 units, and a standard normal base distribution. This choice provides a sufficiently flexible function capable of approximating complex posterior distributions.

(4) **Training procedure.** The parametrization of our normalizing flow is obtained by minimizing the loss function (15) using the ADAM optimizer (Kingma and Ba, 2017) with default parameters, a learning rate of $5 \cdot 10^{-4}$ and a batch size of 100. Except for a few validation experiments, we have used $N = 10^5$ simulated data points to approximate the posterior distribution.

4.2 Simulated dMRI data

We first validate our proposed method using simulated dMRI data. For this, we fix a parameter vector θ_0 based on a plausible biophysical configuration (Palombo et al.) and generate a simulated observation \mathbf{x}_0 associated to it. Our goal, then, is to check whether the posterior distribution $p(\theta|\mathbf{x}_0)$ concentrates around the ground truth parameter, i.e. if it is peaked around the true values of the parameters in θ_0 . If this is the case, we can assert that the LFI procedure is capable of inverting our non-linear model successfully.

The simplest way of generating an observation from the ground truth parameter would be to use the forward model defined in Section 4.1. However, such validation could yield unrealistically good results, since the posterior approximation is trained on data points generated in the same way. We have decided, therefore, to use a different approach in which the dMRI signals are simulated following a setup which is closer to what we would expect from real experimental experiments. This is based on two steps. Firstly, we use the `dmipy` simulator (Fick et al., 2019) to simulate the three-compartment model described in Section 3.1 and obtain a dMRI signal \mathbf{S}_0 . Then, we calculate the summary statistics of this signal as defined in Section 3.2 to reduce the dimensionality of the observation and obtain a feature vector \mathbf{x}_0 .

We have carried out our simulations on `dmipy` considering three different kinds of acquisition setup. They all have b -shells with 128 uniformly distributed directions, but they differ in their b -values and acquisition times:

- Setup `Ideal` corresponds to a rather “comfortable” case with 10 b -values between 0 and $10 \text{ ms } \mu\text{m}^{-2}$. We use $\delta/\Delta = 12.9/21.8 \text{ ms}$ as in the HCP MGH database.
- Setup `HCP MGH` reproduces the setup from the HCP MGH dataset, with 5 b -values: 0, 1, 3, 5, and $10 \text{ ms } \mu\text{m}^{-2}$ and $\delta/\Delta = 12.9/21.8 \text{ ms}$. Since the Spiked LEMONADE approximation (3.2) requires at least three b -values inferior to $2.5 \text{ ms } \mu\text{m}^{-2}$, we extrapolated an extra b -shell at $0.1 \text{ ms } \mu\text{m}^{-2}$ using MAPL (Fick et al., 2016).
- Setup `HCP 1200` reproduces the setup from the HCP 1200 dataset, with only 3 b -values: 1, 2, and $3 \text{ ms } \mu\text{m}^{-2}$ and $\delta/\Delta = 10.6/43.1 \text{ ms}$. An extra b -shell at $2.8 \text{ ms } \mu\text{m}^{-2}$ has been interpolated to be used in the RTOP approximation.

Note that in the simulations with all setups we have used the three b -shells with the lowest b -values for the small q -value approximation (Spiked LEMONADE), and the three largest q -values for the RTOP approximation.

334 4.3 HCP MGH dataset

335 After validating our proposal on different simulated settings, we carried out our analysis
 336 on two publicly available databases. Our goal was to estimate the brain tissue parameters
 337 for each voxel in the dMRI acquisitions and determine how these parameters vary across
 338 the human brain grey matter. Because of the probabilistic framework that we use, these
 339 estimates are accompanied of credible intervals that can be used to inform our degree of
 340 confidence of these estimates.

341 Our first analysis was on the HCP MGH Adult Diffusion database (Setsompop et al.,
 342 2013). This database is composed of 35 subjects with $\delta/\Delta = 12.9/21.8$ ms and $b = 1, 3, 5,$
 343 10 ms μm^{-2} . We used the 3, 5 and 10 ms μm^{-2} b -values for the RTOP approximation (i.e.
 344 the large q -value analysis). For the Spiked LEMONADE approximation, we fitted MAPL on
 345 the 0, 1 and 3 ms μm^{-2} b -values to reduce noise, and interpolated a b -value of 0.1 ms μm^{-2} to
 346 improve the estimations. D_e was estimated as 1/3 of the mean diffusivity in the ventricles.

347 The spatial distribution of the estimated parameters were mapped to the MNI tem-
 348 plate, averaged over all subjects, and then projected onto an inflated cortical surface using
 349 `FreeSurfer`. We have then evaluated whether such distributions seemed physiologically
 350 plausible by using the Brodmann atlas (Brodmann, 1909; Zilles, 2018), which is a parcel-
 351 lation of the brain based on cytoarchitecture features. In addition, we compared qualita-
 352 tively the results of soma estimations with nissl-stained histological images of cytoarchitec-
 353 ture (Allman et al., 2010; Amunts et al., 1999; Geyer et al., 1999).

354 4.4 HCP 1200

355 We proceeded with our analysis on real data using a more challenging database, in which
 356 the dMRI signals were acquired with only a few b -values. Our goal was to demonstrate that
 357 the credible regions obtained via the posterior approximation can be used to inform which
 358 parameters remain possible to estimate even in very challenging situations. Note that this
 359 unlocks the door to the analysis of any dMRI database, since the estimates always come
 360 with a “quality certificate”.

361 We applied our pipeline to a subset of the HCP 1200 database. We randomly picked
 362 30 subjects, to have an identical number of subjects to that in our analysis of the HCP
 363 MGH database. The data were acquired for b -values equal to 1, 2 and 3 ms μm^{-2} , with
 364 $\delta/\Delta = 10.6/43.1$ ms. Using MAPL, we interpolated a b -shell at 2.8 ms μm^{-2} to improve
 365 the computation of the summary statistics. We used all the three lowest b -values for the
 366 Spiked LEMONADE approximation, and $b = 2, 2.8$ and 3 ms μm^{-2} for the high b -value
 367 approximation based on RTOP. Similarly to the HCP MGH dataset, we have averaged
 368 the parameter estimations in a common space (MNI) and then projected the resulting
 369 parametric maps onto an inflated cortical surface.

370 5. Results

371 In this section, we describe our results obtained on simulated data and two real datasets.
 372 All experiments have been implemented with Python (Python Software Foundation, 2017)
 373 using several scientific packages: dMRI signals were simulated with the package `dmipy` (Fick
 374 et al., 2019) and processed using `dipy` (Garyfallidis et al., 2014) or custom implementations

375 based on `numpy` (Harris et al., 2020). We used the `sbi` (Tejero-Cantero et al., 2020) and
 376 `nflows` (Durkan et al., 2020) packages for carrying out the LFI procedures and combined
 377 them with data structures and functions from `pyTorch` (Paszke et al., 2019). The fig-
 378 ures of results on real experimental data were generated with `mayavi` (Ramachandran and
 379 Varoquaux, 2011).

380 5.1 Simulated data

381 **Validating the LFI pipeline.** In this first round of experiments, our aim was to check
 382 whether the LFI pipeline worked correctly on a setting where we knew the true values of
 383 the parameter θ_0 (ground truth) generating the observed data \mathbf{x}_0 . Furthermore, we wanted
 384 to confirm whether the use of summary statistics for the dMRI signal actually conveyed any
 385 improvements in the parameter estimation. We have considered, therefore, three different
 386 cases:

- 387 – **Case 1.** Generate \mathbf{x}_0 directly from θ_0 using the forward model defined in Section 4.1.
 388 This is a rather favorable case for our posterior approximation, since it is applied on
 389 a data point generated in the same way as the dataset in which it was trained.
- 390 – **Case 2.** Generate a dMRI signal \mathbf{S}_0 from θ_0 using `dmipy` and then obtain \mathbf{x}_0 by
 391 calculating the summary statistics presented in Section 3.2. We use the same posterior
 392 approximator from **Case 1**, meaning that the data point in inference time is generated
 393 differently from those for the training procedure. This is the actual realistic case that
 394 interests us the most.
- 395 – **Case 3.** Generate a dMRI signal \mathbf{S}_0 from θ_0 using `dmipy` and do not use any summary
 396 statistics for the model inversion, i.e. consider $\mathbf{x}_0 = \mathbf{S}_0$. Note that the posterior
 397 approximator has to be trained on a dataset with observations consisting of dMRI
 398 signals, so it is different from the approximators in **Case 1** and **Case 2**. Depending
 399 on how the LFI pipeline behaves on this case, the use of summary statistics can be
 400 justified or not.

All simulations were carried out with setup `Ideal`, which corresponds to an ideal dMRI acquisition scheme, and the posterior approximators were trained with $N = 10^5$ simulated data points. While we have validated the LFI pipeline on multiple choices of physiologically relevant ground truth parameters θ_0 , Figure 1 portrays the results only for

$$\theta_0 = (D_n, C_s, p_2, f_s, f_n, f_{ECS}) = (2.5 \mu\text{m ms}^{-2}, 617 \mu\text{m}^2, 0.50, 0.15, 0.45, 0.40) .$$

401 This choice of parameters represent a sample tissue containing pyramidal neurons of radius
 402 $12 \mu\text{m}$ (Palombo et al.) and diffusivity $3 \mu\text{m ms}^{-2}$. The three compartment proportions
 403 were chosen in accordance to reported values from histology of human grey matter tis-
 404 sue (Shapson-Coe et al., 2021). Results are presented in Figure 1. We see that in **Case**
 405 **1** the marginals of the posterior distribution are well concentrated around the values of
 406 the ground truth parameter θ_0 . This confirms that the posterior approximator successfully
 407 inverts the non-linear model using the examples in the training set. We also see that the
 408 parameter estimation in **Case 2** captures very well most of the true values of the parame-
 409 ters, indicating both that our posterior approximator is robust to observed data generated

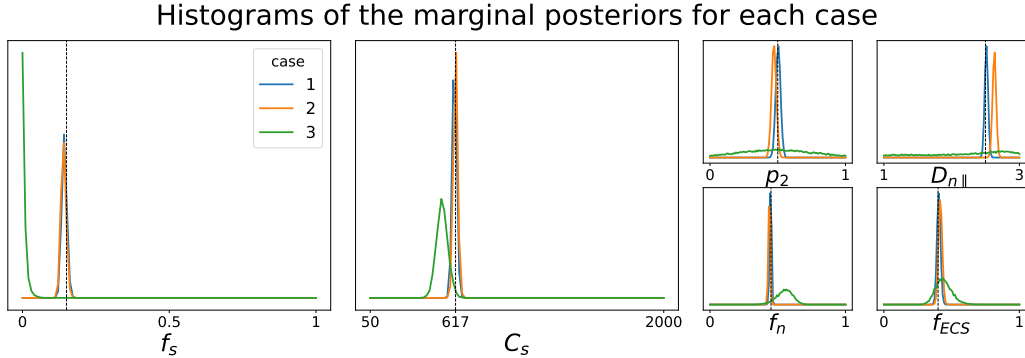


Figure 1: Histograms of 10^4 samples of the approximate posterior distribution in three different base cases (see text for details). Vertical black dashed lines represent ground truth values of θ_0 which generated the observed dMRI signals. We observe that the marginals tend to concentrate around the ground truth parameters when the observed summary statistics are obtained directly from the parameters (**Case 1**) and have a small bias when the signals are generated using `dmipy` (**Case 2**). The figure also shows that inverting the model directly from the dMRI signals leads to rather poor results (**Case 3**).

410 differently from its training set and that our summary statistics are descriptive enough to
 411 synthesize each tissue configuration. Finally, the poor results in the estimation for **Case 3**
 412 reflect the largely indeterminate system of equations that results from trying to infer the
 413 tissue parameters directly from the dMRI signals. This behavior was expected, as similar
 414 degeneracies were shown for a simpler model in Novikov et al. (2018a).

415 **Influence of the number of b -values.** We have also investigated how different choices
 416 of b -values in the acquisition scheme affect the quality of the parameter estimation using
 417 the posterior approximation. Note that these choices have no influence on how the posterior
 418 approximator is obtained, since it is trained on data generated directly from the relations
 419 between tissue parameters and the diffusion summary statistics, in which the b -values do
 420 not interact. In fact, the different acquisition setups that we consider have only an impact
 421 over the observed data point generated via `dmipy`. Figure 2 portrays the marginal posterior
 422 distributions for each tissue parameter in setups **Ideal**, **HCP MGH**, and **HCP 1200**. We see
 423 that estimations in the **Ideal** setup (equivalent to Case 2 in Figure 1) are very much
 424 concentrated around the true values of the parameters. Results with **HCP MGH** present a
 425 small bias for most parameters but capture well the soma parameters (f_s and C_s), whereas
 426 the results with setup **HCP 1200** present a larger bias for most parameters. They nevertheless
 427 remain reliable for the estimation of soma parameters. Note that the **Ideal** and **HCP MGH**
 428 ground truth value of C_s (vertical black dashed line) is different from the one of the **HCP**
 429 **1200** setup (vertical green dashed line), because of their different diffusion times. These
 430 simulations allow us to have a fair confidence in the estimations on real data, presented in
 431 Sections 5.2 and 5.3.

432 **Our new parameter avoids model indeterminacy.** In Section 3.1, we introduce the
 433 parameter C_s , which serves as a proxy of the soma radius and provides key information on

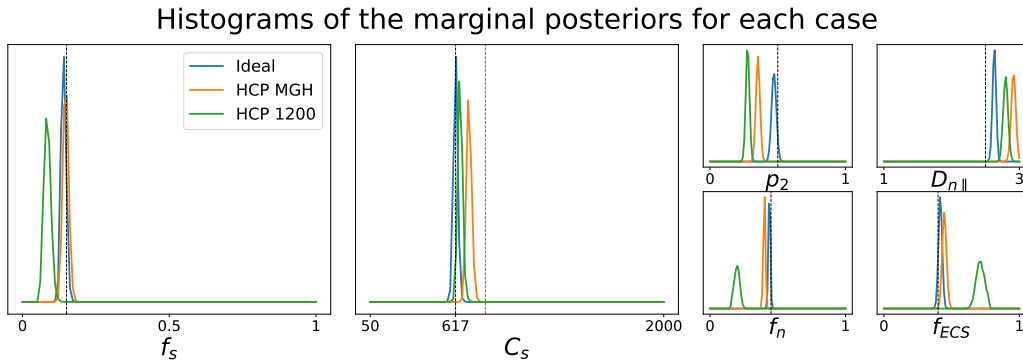


Figure 2: Histograms of 10^4 samples of the approximate posterior distribution with observed dMRI signals generated under three acquisition setups (see text for details). Vertical black dashed lines represent ground truth values of θ_0 which generated the observed signals (the green dashed line indicates the C_s value expected for the HCP 1200 setup). We see that while the **Ideal** case delivers very good estimates, the results for the two other setups are only reliable for a subset of the tissue parameters.

434 the soma compartment. In this experiment, we illustrate the results of our model inversion
 435 if we had not defined parameter C_s .

436 Figure 3 presents the marginal posterior distributions of r_s and D_s as well as their joint
 437 distribution using the setup **Ideal**. To obtain these results, we have altered our LFI pipeline
 438 so to consider an extended parameter vector including r_s and D_s . The prior distribution
 439 reflects our assumption that $r_s \in [10^{-5}, 30] \mu\text{m}$ and $D_s \in [1, 3] \mu\text{m ms}^{-2}$ and we consider
 440 ground truth parameters $r_s = 17.5 \mu\text{m}$ and $D_s = 2.3 \mu\text{m ms}^{-2}$. We note that in addition to
 441 larger marginal posterior distributions for each parameter, the joint posterior has a valley
 442 of large values for the (r_s, D_s) pair, including the ground truth parameters. This result is
 443 typical of non-injective models, i.e. models for which several input parameters may yield
 444 the same output observation, and is an important asset of a probabilistic framework such
 445 as ours.

446 **Assessing the variances of estimated parameters.** Deriving the posterior distri-
 447 butions of the parameter vectors allows us to report the values of the most likely tissue
 448 parameters for a given observation, along with our certitude regarding our inference. Fig-
 449 ure 4 presents the logarithm of the standard deviation of the marginal posterior samples
 450 for different ground truth parameter choices (varying f_s and f_n) under setup **Ideal**. These
 451 values indicate how sharp a posterior distribution is and, therefore, quantify the quality of
 452 the fit. We observe larger standard deviations in the absence (or weak presence) of soma
 453 compartments in the mixture signal, e.g., the standard deviation of C_s is large when few or
 454 no somas are present ($f_s \approx 0$). This is to be expected, since the lack of contribution from
 455 the somas in the diffusion signal makes it difficult to estimate parameters related to them.

456 5.2 HCP MGH

457 Our results on simulated data show that, although we manage to invert very well the brain
 458 tissue parameter on settings for which the dMRI signal is obtained with several b -values,

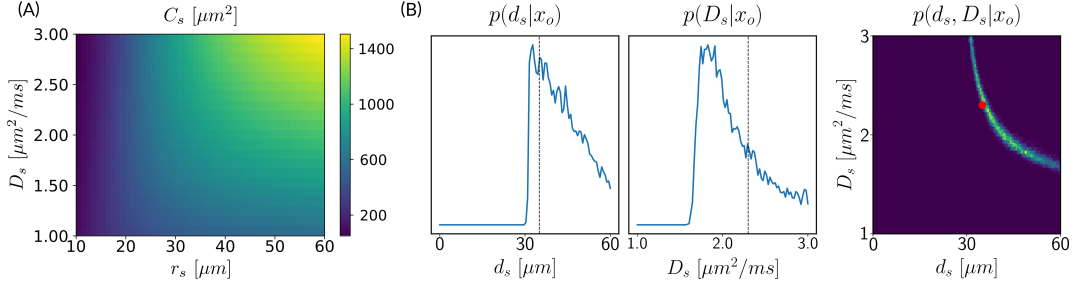


Figure 3: (A) C_s dependence on soma radius r_s and diffusivity D_s . We see that there are several values of (d_s, D_s) that yield the same C_s . (B) and (C) Histograms of 10^4 samples from the marginal and joint posterior distributions of $d_s = 2r_s$ and D_s . The ridge in the joint distribution indicates that there are several possible values for the pair (d_s, D_s) with high probability, which are those yielding the same C_s . Estimating C_s directly bypasses this indeterminacy.

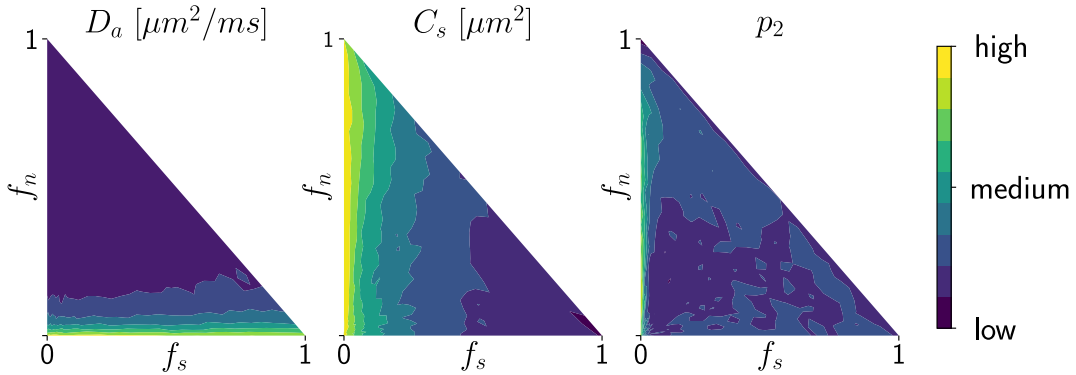


Figure 4: Logarithm of the standard deviations for the marginal posterior distribution of D_n , C_s , and p_2 with different choices of ground truth parameters (varying f_s and f_n). Since the ranges of values for each plot were quite different, we labeled the colorbar in terms of {‘high’, ‘medium’, ‘low’} values of standard deviations so to provide mainly qualitative information to the reader. We see that when the signal fraction of somas decreases ($f_s \rightarrow 0$) the standard deviation of the C_s -estimation increases; and when less neurites are present ($f_n \rightarrow 0$) the standard deviation of p_2 and D_n increase.

459 the estimates for more realistic settings are less robust and demand a more subtle analysis.
 460 Indeed, we have observed that for both setups HCP MGH and HCP 1200 the soma parameters
 461 seem to be rather well estimated without too much bias, which has lead us to consider
 462 mainly these parameters in our interpretations of the results.

463 Figure 5 presents the results on the HCP MGH dataset. We have masked our results so to
 464 show only areas where parameters were deemed stable, i.e. when the values were larger than
 465 2 times the LFI-obtained standard deviations of the fitted posterior. We observe, therefore,
 466 a lack of stability on small sections including the auditory cortex and the precentral gyrus
 467 fundus. Our figure assesses qualitatively the results on soma size by comparing with nissl-
 468 stained histological studies (Allman et al., 2010; Amunts et al., 1999; Geyer et al., 1999).
 469 Our comparison shows good agreement between different cortical areas and the parameter
 470 C_s , which, under nearly-constant intra-soma diffusion D_s , is modulated by soma size.

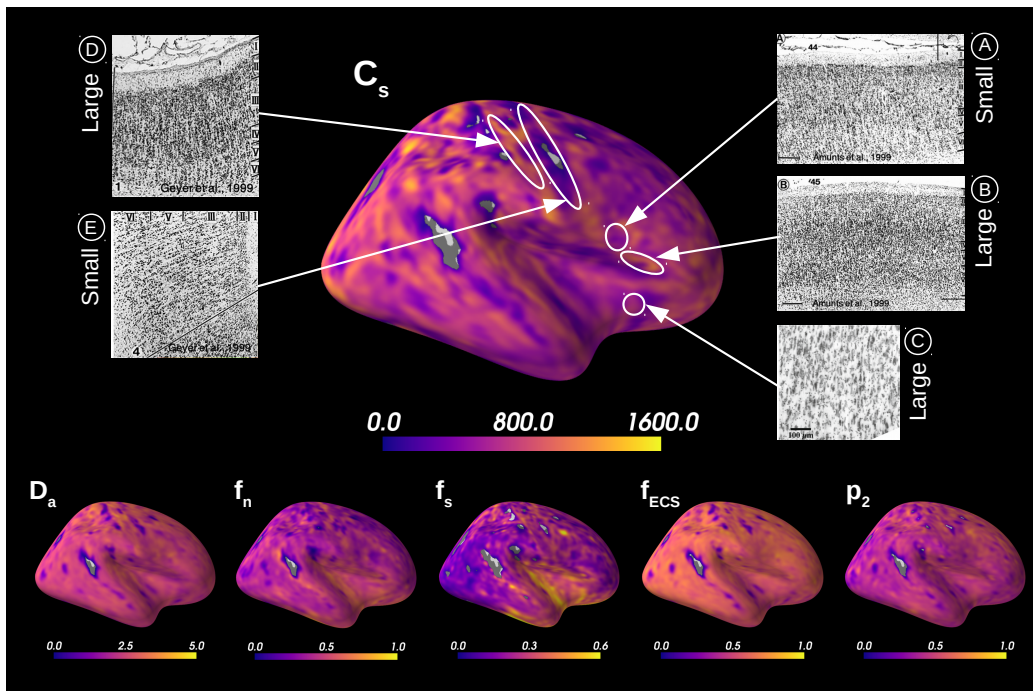


Figure 5: Microstructural measurements averaged over 31 HCP MGH subjects. We deemed stable measurements with a z-score larger than 2, where the standard deviation on the posterior estimates was estimated through our LFI fitting approach. In comparing with Nissl-stained cytoarchitectural studies we can qualitatively evaluate our parameter C_s : Broadmann area 44 (A) has smaller soma size in average than area 45 (B) (Amunts et al., 1999); large von Economo neurons predominate the superior anterior insula (C) (Allman et al., 2010); precentral gyrus (E) shows very small somas while post-central (D) larger ones (Geyer et al., 1999).

471 Interestingly, most regions of Figure 6 in which the parameter estimation has low con-
 472 fidence are located in the fundus of the sulci. Two main hypothesis could explain such be-
 473 havior. Firstly, brain regions which are very curved may be more prone to mixing between
 474 tissue layers and CSF, which generates noisier signals. Thus, the estimation of summary

475 statistics becomes more difficult and the posterior distributions tend to be wider. Another
 476 possible explanation, based on cytoarchitecture considerations, points out the fact that the
 477 fundus of sulci is where sharp changes in cellular populations happen (Brodmann, 1909;
 478 Pandya et al., 2015). A mixing of several types of neurons in one voxel could generate
 479 multi-modal posterior distributions, and hence a region with large variance.

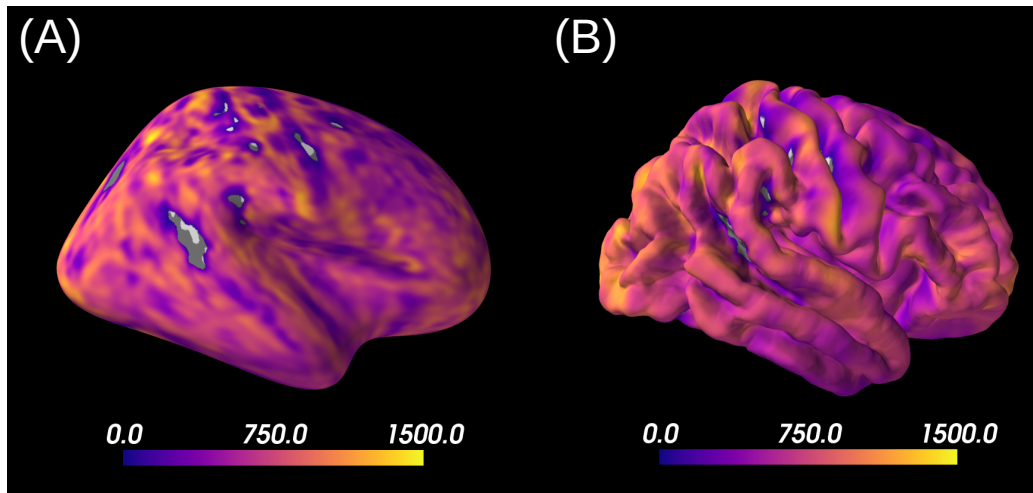


Figure 6: C_s estimations averaged over 31 HCP MGH subjects, with unstable results masked, projected onto inflated (A) and pial (B) surfaces. Interestingly, low confidence areas correspond to the fundus of the sulci.

480 Figure 7A reports the soma proportion (parameter f_s) averaged over 31 HCP MGH
 481 subjects, masking unstable results. Mean soma signal proportion in the cortex equals
 482 0.32 ± 0.22 (mean computed in trusted estimations only). These results are coherent with
 483 the mean volume occupancy of 10 – 20% observed in grey matter (Shapson-Coe et al.,
 484 2021). To interpret the results at a region-based level, we have superimposed the soma
 485 proportion estimations with the Brodmann atlas. We observe a general agreement between
 486 the estimations and the atlas, and more particularly in the somatosensory and Broca’s
 487 areas. A clear difference in soma proportion can be observed in the 12 regions, as presented
 488 in the barplot.

489 Despite the C_s parameter being useful for avoiding indeterminacies in the model in-
 490 version, its biological interpretation remains difficult. With the goal of relating our C_s
 491 estimation with physiological insight, we estimated soma radius r_s by fixing soma diffu-
 492 sivity. Similarly to the SANDI method, we fixed soma diffusivity D_s to the value of the
 493 self-diffusion coefficient of free water ($3 \mu\text{m ms}^{-2}$). Note that this value could be adjusted
 494 for each voxel, and is only used here in a matter of comparison and interpretation. Us-
 495 ing a fixed-point method, we computed the soma radius r_s from the averaged C_s map
 496 obtained from our posterior distribution. The results are portrayed in Figure 7B. We see
 497 that the estimated soma radius vary between 8 and $14 \mu\text{m}$, which is in accordance with
 498 histology (Palombo et al.). Mean C_s values are presented in the barplot beneath the soma
 499 radius estimations in the different Brodmann regions.

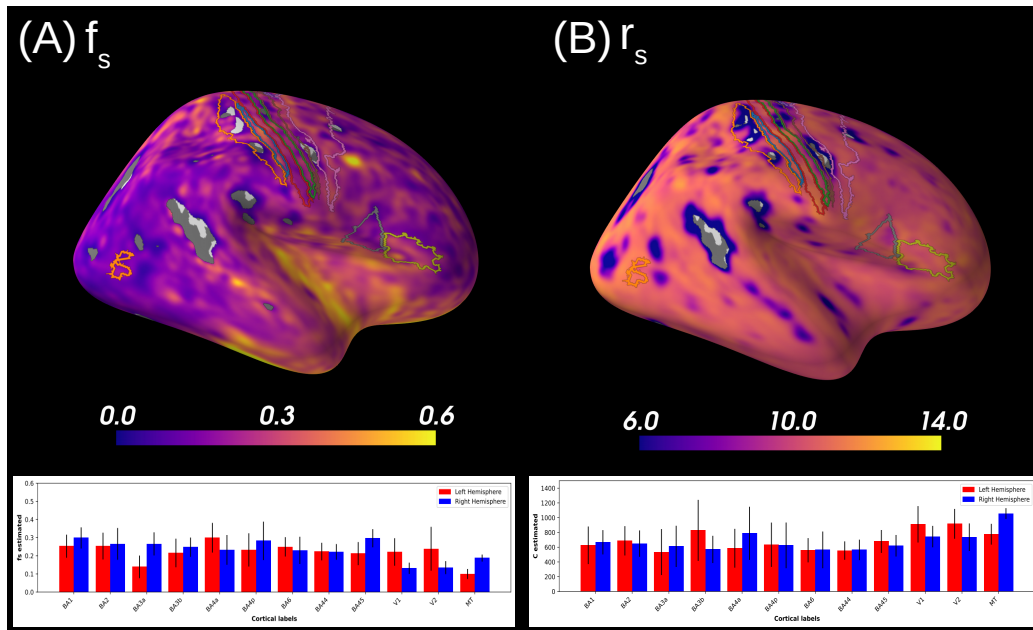


Figure 7: (Left) The average soma proportion (f_s) over the 31 HCP MGH subjects is projected onto an inflated cortical surface, with unstable results masked; see text for details. We also show the main Brodmann areas available on FreeSurfer. The mean values of f_s over these regions (using only reliable estimations) are reported on the bar plots below. (Right) Soma radius map computed from C_s with soma diffusivity fixed to $D_s = 3 \mu\text{m ms}^{-2}$, averaged on 31 subjects. The bar plots below report the C_s mean values on main Brodmann areas. We observe a good agreement between our reported values and the Brodmann areas.

500 5.3 HCP 1200

501 Figure 8 shows the results obtained on a database with only three b -shells. We see that
 502 55 % of the C_s estimations on brain grey matter is considered unstable and is, therefore,
 503 masked. Indeed, our q -bounded RTOP approximation relies on high b -values, where the
 504 signal is expected to have converged towards a value that depends on the radius of the
 505 soma. The larger the soma, the sooner the q -bounded RTOP converges. However, the
 506 largest b -value contained in this database equals $3 \text{ms } \mu\text{m}^{-2}$, which is not enough for the
 507 signal to have converged. The poor quality of the summary statistics estimation leads to
 508 rather wide posterior distributions, resulting in unreliable results, as shown in simulations.
 509 Note, however, that estimations of the superior temporal gyrus for example are not masked,
 510 and both data sets seem to indicate large neurones in that region. The estimation of D_a
 511 is considered as unstable for 98.7% of the data set. This behavior was expected, given the
 512 results presented in Figure 2, for which a large bias is observed.

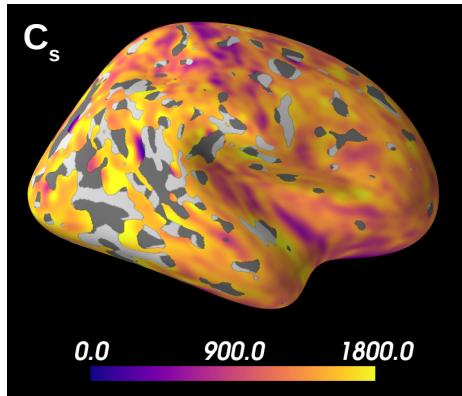


Figure 8: Spatial distribution of the C_s estimations averaged over 30 HCP 1200 subjects and projected onto an inflated cortical surface, with unstable estimations masked. Due to the scarce and rather low b -values ($\leq 3 \text{ ms } \mu\text{m}^{-2}$) used in the database, the results are very unstable and, therefore, many voxels are discarded.

513 6. Discussion

514 6.1 Validation simulated data

515 An important aspect of our work is the thorough validation that we have carried out on
 516 simulated data, using different acquisition setups and ways of generating diffusion signals.
 517 Part of this validation had the goal of ensuring that our method gave consistent results
 518 on simple standard cases, as confirmed by the results in Figure 1. Additionally, we have
 519 demonstrated the benefits of using summary statistics for describing dMRI signals, attaining
 520 better parameter estimates when using them instead of directly manipulating the diffusion
 521 signals. Such results are very encouraging and should push other researchers into using
 522 these summary statistics for processing their dMRI signals.

523 Another relevant byproduct of our validation was observing that the quality of the pa-
 524 rameter estimations depends heavily on the distribution of b -values used to acquire and
 525 simulate the dMRI signals. Indeed, if only small b -values are available, the summary statis-
 526 tics of dMRI signals are poorly estimated and the parameter estimation too. Based on this
 527 observation, we were able to apply our LFI pipeline to real datasets HCP MGH and HCP
 528 1200 knowing in advance the limitations of our methodology; for example, we knew from
 529 which parameters we could expect good estimates (mostly soma-related ones) and which
 530 ones should not be taken into account in our analysis.

531 6.2 C_s : A proxy to soma size

532 Estimating both soma radius (r_s) and diffusivity (D_s) with diffusion MRI is a challenging
 533 task. When trying to estimate them separately, we can expect a ‘banana-shape’ in their
 534 joint posterior distribution as shown in Figure 3. This indicates that several values of the
 535 pair (r_s, D_s) can explain the observed signal with high probability and, therefore, one is
 536 confronted with model indeterminacy.

537 The new parameter C_s that we introduce in this paper is modulated both by the soma
 538 radius and its diffusivity. Thus, estimating it directly avoids problems of indeterminacy, as
 539 shown in Figure 1, for example. Note, however, that avoiding such indeterminacy comes
 540 with the price of losing specificity and, therefore, physiological interpretations. Fortunately,
 541 acquisitions in the narrow or wide pulse regimes allow us to better interpret estimations of
 542 C_s , as it only depends on r_s or D_s (see Section 3.1).

543 6.3 Can I apply this approach to my data?

544 One of the main benefits of a probabilistic framework is that it can be applied to any data
 545 set or acquisition setup without too much hesitation, since the estimates always come with
 546 a “quality certificate” described by the credible regions derived from the posterior distri-
 547 bution. We have benefited from this feature when creating all figures related to databases
 548 HCP MGH and HCP 1200, since they allow us to mask regions for which the variance of the
 549 parameter estimation is too high. We can also use it as a proxy to identify regions for which
 550 the three-compartment model is adequate or not, or assessing whether the distribution of
 551 b -values used to acquire the observed diffusion signal is sufficiently informative.

552 6.4 Limitations and perspectives

553 There are many extensions that we could envision for the method that we propose.

554 For instance, the three-compartment model that we use approximates well the intra-
 555 cellular signal in brain grey matter tissue by adding a sphere compartment to account
 556 for soma presence (Palombo et al., 2020). However, this is a rather simplified model and
 557 could be improved; for instance, the geometry of ECS is very restricted and tortuous, and
 558 diffusion signals have been proven to deviate from a mono-exponential behavior (Vincent
 559 et al., 2021). A first improvement could be, therefore, to deviate from modelling the diffusion
 560 with a simple isotropic Gaussian and consider more complex geometric representations.

561 Also, soma-related parameters are better estimated in tissues where there are many
 562 somas. Consequently, the method that we present is more suitable to grey matter than for
 563 white matter. However, applying a LFI pipeline on a modified model in which the contribu-
 564 tions of the somas are discarded would improve the estimates of tissue parameters for white
 565 matter. This new pipeline would then be seen as a new way to solve the LEMONADE
 566 system of equations proposed by Novikov et al. (2018b) within a probabilistic framework.

567 Finally, we could leverage the modularity of the probabilistic framework and apply our
 568 methodology to other kinds of models in medical imaging. This could be done by simply
 569 re-defining the forward model and the summary statistics defined in Section 4.1. Such
 570 flexibility confirms the generality of our approach and the potential impact that it might
 571 have in the medical imaging field.

572 7. Conclusion

573 Quantifying grey matter tissue composition is challenging. In this work, we have presented
 574 a methodology to estimate the parameters of a model that best fit an observed data point,
 575 and also their full posterior distribution. This rich description provides many useful tools,
 576 such as assessing the quality of the parameter estimation or characterizing regions in the

577 parameter space where it is harder to invert the model. The inclusion of such “quality
 578 certificate” accompanying our parameter estimation is very useful in practice and allows
 579 one to apply the pipeline on any kind of database and know to which degree one can
 580 trust the results. Moreover, our proposal alleviates limitations from current methods in
 581 the literature by not requiring physiologically unrealistic constraints on the parameters and
 582 avoiding indeterminacies when estimating them.

583 To conclude, we believe that our approach based on Bayesian inference with modern
 584 tools from neural networks is a promising one that can easily be applied to other applications
 585 in the medical imaging field: one only needs to define a sufficiently rich model describing a
 586 certain phenomenon of interest and the LFI pipeline will manage to invert it and provide
 587 a related posterior distribution. We expect, therefore, that other researchers will find this
 588 contribution valuable for their own applications and see such probabilistic approach more
 589 often used in the literature.

590 Acknowledgments

591 Experiments on diffusion MRI data were made possible thanks to DiPy (Garyfallidis et al.,
 592 2014), as well as the scientific Python ecosystem: Python (Python Software Foundation,
 593 2017), Matplotlib (Hunter, 2007), Numpy (Harris et al., 2020), Scipy (Virtanen et al.,
 594 2020), Mayavi (Ramachandran and Varoquaux, 2011), SBI (Tejero-Cantero et al., 2020)
 595 and PyTorch (Paszke et al., 2019).

596 This work was supported by grants ERC-StG NeuroLang ID:757672 and ANR-NSF
 597 NeuroRef to DW and the ANR BrAIN (ANR-20-CHIA-0016) grant to AG.

598 Ethical Standards

599 The work follows appropriate ethical standards in conducting research and writing the
 600 manuscript, following all applicable laws and regulations regarding treatment of animals or
 601 human subjects.

602 Conflicts of Interest

603 We declare we don’t have conflicts of interest.

604 References

- 605 John M. Allman, Nicole A. Tetreault, Atiya Y. Hakeem, Kebreten F. Manaye, Katerina Semendeferi,
 606 Joseph M. Erwin, Soyoung Park, Virginie Goubert, and Patrick R. Hof. The von Economo neurons
 607 in frontoinsular and anterior cingulate cortex in great apes and humans. Brain Structure and
 608 Function, 214:495–517, June 2010.
- 609 Katrin Amunts, Axel Schleicher, Uli Bürgel, Hartmut Mohlberg, Harry B.M. Uylings, and Karl Zilles.
 610 Broca’s region revisited: Cytoarchitecture and intersubject variability. Journal of Comparative
 611 Neurology, 412(2):319–341, 1999.

- 612 Balin Balinov, Jönsson, Per Linse, and Olle Söderman. The NMR Self-Diffusion Method Applied
613 to Restricted Diffusion. Simulation of Echo Attenuation form Molecules in Spheres and between
614 Planes. Journal of Magnetic Resonance, Series A, 104(1):17–25, 1993.
- 615 Korbinian Brodmann. Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien
616 dargestellt auf Grund des Zellenbaues. Barth, 1909.
- 617 Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference.
618 Proceedings of the National Academy of Sciences, 117(48):30055–30062, 2020.
- 619 Gwenaëlle Douaud, Ricarda A. L. Menke, Achim Gass, Andreas U. Monsch, Anil Rao, Brandon
620 Whitcher, Giovanna Zamboni, Paul M. Matthews, Marc Sollberger, and Stephen Smith. Brain
621 microstructure reveals early abnormalities more than two years prior to clinical progression from
622 mild cognitive impairment to alzheimer’s disease. Journal of Neuroscience, 33(5):2147–2155, 2013.
623 ISSN 0270-6474. doi: 10.1523/JNEUROSCI.4437-12.2013. URL [https://www.jneurosci.org/
624 content/33/5/2147](https://www.jneurosci.org/content/33/5/2147).
- 625 Conor Durkan, Artur Bekasov, Iain Murray, and George Papamakarios. nflows: normalizing flows
626 in PyTorch. November 2020. doi: 10.5281/zenodo.4296287.
- 627 Rutger H. J. Fick, Demian Wassermann, and Rachid Deriche. The Dmipy Toolbox: Diffusion
628 MRI Multi-Compartment Modeling and Microstructure Recovery Made Easy. Frontiers in
629 Neuroinformatics, 13:64, October 2019.
- 630 Rutger H.J. Fick, Demian Wassermann, Emmanuel Caruyer, and Rachid Deriche. MAPL: Tissue
631 microstructure estimation using Laplacian-regularized MAP-MRI and its application to HCP
632 data. NeuroImage, 134:365–385, July 2016.
- 633 Eleftherios Garyfallidis, Matthew Brett, Bagrat Amirbekian, Ariel Rokem, Stefan Van Der Walt,
634 Maxime Descoteaux, and Ian Nimmo-Smith. Dipy, a library for the analysis of diffusion mri data.
635 Frontiers in Neuroinformatics, 8:8, 2014. ISSN 1662-5196. doi: 10.3389/fninf.2014.00008. URL
636 <https://www.frontiersin.org/article/10.3389/fninf.2014.00008>.
- 637 A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. Bayesian Data
638 Analysis, Third Edition. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis,
639 2013. ISBN 9781439840955. URL <https://books.google.fr/books?id=ZXL6AQAAQBAJ>.
- 640 Mathieu Germain, Karol Gregor, Iain Murray, and Hugo Larochelle. Made: Masked autoencoder
641 for distribution estimation. In Proceedings of the 32nd International Conference on Machine
642 Learning, volume 37, pages 881–889. PMLR, 2015.
- 643 Stefan Geyer, Axel Schleicher, and Karl Zilles. Areas 3a, 3b, and 1 of Human Primary Somatosensory
644 Cortex. NeuroImage, 10(1):63–83, July 1999. ISSN 10538119.
- 645 Pedro J Gonçalves, Jan-Matthis Lueckmann, Michael Deistler, Marcel Nonnenmacher, Kaan Öcal,
646 Giacomo Bassetto, Chaitanya Chintaluri, William F Podlaski, Sara A Haddad, Tim P Vogels,
647 David S Greenberg, and Jakob H Macke. Training deep neural density estimators to identify
648 mechanistic models of neural dynamics. eLife, 9, September 2020.
- 649 Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016. [http:
650 //www.deeplearningbook.org](http://www.deeplearningbook.org).
- 651 Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair,
652 Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.

- 653 David Greenberg, Marcel Nonnenmacher, and Jakob Macke. Automatic posterior transformation
 654 for likelihood-free inference. In Proceedings of the 36th International Conference on Machine
 655 Learning, volume 97, pages 2404–2414. PMLR, 2019.
- 656 Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David
 657 Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti
 658 Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández
 659 del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy,
 660 Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array program-
 661 ming with NumPy. Nature, 585(7825):357–362, September 2020. doi: 10.1038/s41586-020-2649-2.
 662 URL <https://doi.org/10.1038/s41586-020-2649-2>.
- 663 John D Hunter. Matplotlib: A 2d graphics environment. Computing in science & engineering, 9(3):
 664 90–95, 2007.
- 665 Ileana O. Jelescu and Matthew D. Budde. Design and validation of diffusion MRI models of white
 666 matter. Frontiers in Physics, 5, November 2017. doi: 10.3389/fphy.2017.00061. URL <https://doi.org/10.3389/fphy.2017.00061>.
- 668 Ileana O. Jelescu, Jelle Veraart, Els Fieremans, and Dmitry S. Novikov. Degeneracy in model
 669 parameter estimation for multi-compartmental diffusion in neuronal tissue: Degeneracy in Model
 670 Parameter Estimation of Diffusion in Neural Tissue. NMR in Biomedicine, 29(1):33–47, January
 671 2016.
- 672 Ileana O. Jelescu, Alexandre de Skowronski, Marco Palombo, and Dmitry S. Novikov. Neurite
 673 exchange imaging (nexi): A minimal model of diffusion in gray matter with inter-compartment
 674 water exchange, 2021.
- 675 Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.
- 676 Diederik P Kingma and Max Welling. Auto-encoding variational bayes, 2014.
- 677 Hua Li, Xiaoyu Jiang, Jingping Xie, John C. Gore, and Junzhong Xu. Impact of transcytolemmal
 678 water exchange on estimates of tissue microstructural properties derived from diffusion MRI.
 679 Magnetic Resonance in Medicine, 77(6):2239–2249, June 2016. doi: 10.1002/mrm.26309. URL
 680 <https://doi.org/10.1002/mrm.26309>.
- 681 Vinod Menon, Guillermo Gallardo, Mark A Pinsk, Van-Dang Nguyen, Jing-Rebecca Li, Weidong
 682 Cai, and Demian Wassermann. Microstructural organization of human insula is linked to its
 683 macrofunctional circuitry and predicts cognitive control. elife, 9:e53470, 2020.
- 684 P. P. Mitra, L. L. Latour, R. L. Kleinberg, and C. H. Sotak. Pulsed-field-gradient NMR measurements
 685 of restricted diffusion and the return-to-origin probability. Journal of Magnetic Resonance, 114:
 686 47–58, 1995.
- 687 J. S. Murday and R. M. Cotts. Self-diffusion coefficient of liquid lithium. The Journal of Chemical
 688 Physics, 48(11):4938–4945, 1968. doi: 10.1063/1.1668160.
- 689 Dmitry S. Novikov, Els Fieremans, Sune N. Jespersen, and Valerij G. Kiselev. Quantifying brain
 690 microstructure with diffusion MRI: Theory and parameter estimation: Brain microstructure with
 691 dMRI: Theory and parameter estimation. NMR in Biomedicine, page e3998, October 2018a. ISSN
 692 09523480.

- 693 Dmitry S. Novikov, Jelle Veraart, Ileana O. Jelescu, and Els Fieremans. Rotationally-invariant
 694 mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI.
 695 NeuroImage, 174:518–538, July 2018b.
- 696 M Palombo, DC Alexander, and H Zhang. Large-scale analysis of brain cell morphometry informs
 697 microstructure modelling of gray matter.
- 698 Marco Palombo, Andrada Ianus, Michele Guerreri, Daniel Nunes, Daniel C. Alexander, Noam
 699 Shemesh, and Hui Zhang. SANDI: A compartment-based model for non-invasive apparent soma
 700 and neurite imaging by diffusion MRI. NeuroImage, 215:116835, 2020.
- 701 Deepak Pandya, Michael Petrides, and Patsy Benny Cipolloni. Cerebral cortex: architecture,
 702 connections, and the dual origin concept. Oxford University Press, 2015.
- 703 George Papamakarios and Iain Murray. Fast ϵ -free inference of simulation models with bayesian
 704 conditional density estimation. In Advances in Neural Information Processing Systems, volume 29,
 705 pages 1028–1036. Curran Associates, Inc., 2016.
- 706 George Papamakarios, Eric T. Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Bal-
 707 aji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. ArXiv,
 708 abs/1912.02762, 2019.
- 709 Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor
 710 Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Ed-
 711 ward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit
 712 Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An Imperative Style, High-
 713 Performance Deep Learning Library. In Advances in Neural Information Processing Systems
 714 (NeurIPS), page 12, Vancouver, BC, Canada, 2019.
- 715 Python Software Foundation. Python Language Reference, version 3.6, 2017.
- 716 P. Ramachandran and G. Varoquaux. Mayavi: 3D Visualization of Scientific Data. Computing in
 717 Science & Engineering, 13(2):40–51, 2011. ISSN 1521-9615.
- 718 K. Setsompop, R. Kimmlingen, E. Eberlein, T. Witzel, J. Cohen-Adad, J.A. McNab, B. Keil, M.D.
 719 Tisdall, P. Hoecht, P. Dietz, S.F. Cauley, V. Tountcheva, V. Matschl, V.H. Lenz, K. Heberlein,
 720 A. Potthast, H. Thein, J. Van Horn, A. Toga, F. Schmitt, D. Lehne, B.R. Rosen, V. Wedeen,
 721 and L.L. Wald. Pushing the limits of in vivo diffusion MRI for the Human Connectome Project.
 722 NeuroImage, 80:220 – 233, 2013.
- 723 Alexander Shapson-Coe, Michał Januszewski, Daniel R. Berger, Art Pope, Yuelong Wu, Tim Blakely,
 724 Richard L. Schalek, Peter Li, Shuohong Wang, Jeremy Maitin-Shepard, Neha Karlupia, Sven
 725 Dorkenwald, Evelina Sjostedt, Laramie Leavitt, Dongil Lee, Luke Bailey, Angerica Fitzmaurice,
 726 Rohin Kar, Benjamin Field, Hank Wu, Julian Wagner-Carena, David Aley, Joanna Lau, Zudi Lin,
 727 Donglai Wei, Hanspeter Pfister, Adi Peleg, Viren Jain, and Jeff W. Lichtman. A connectomic
 728 study of a petascale fragment of human cerebral cortex. bioRxiv, 2021. doi: 10.1101/2021.05.29.
 729 446289. URL <https://www.biorxiv.org/content/early/2021/05/30/2021.05.29.446289>.
- 730 Scott A. Sisson. Handbook of Approximate Bayesian Computation. Chapman and Hall/CRC,
 731 September 2018.
- 732 A. M. Stuart. Inverse problems: A bayesian perspective. Acta Numerica, 19:451–559, 2010. doi:
 733 10.1017/S0962492910000061.

- 734 Chantal M.W. Tax, Filip Szczepankiewicz, Markus Nilsson, and Derek K. Jones. The dot-
735 compartment revealed? diffusion mri with ultra-strong gradients and spherical tensor encoding
736 in the living human brain. *NeuroImage*, 210:116534, 2020.
- 737 Alvaro Tejero-Cantero, Jan Boelts, Michael Deistler, Jan-Matthis Lueckmann, Conor Durkan, Pe-
738 dro J. Gonçalves, David S. Greenberg, and Jakob H. Macke. sbi: A toolkit for simulation-based
739 inference. *Journal of Open Source Software*, 5(52):2505, 2020. doi: 10.21105/joss.02505.
- 740 Jelle Veraart, Els Fieremans, and Dmitry S. Novikov. On the scaling behavior of water diffusion in
741 human brain white matter. *NeuroImage*, 185:379–387, January 2019.
- 742 Jelle Veraart, Daniel Nunes, Umesh Rudrapatna, Els Fieremans, Derek K Jones, Dmitry S Novikov,
743 and Noam Shemesh. Noninvasive quantification of axon radii using diffusion MRI. *eLife*, 9,
744 February 2020.
- 745 Mélissa Vincent, Mylène Gaudin, Covadonga Lucas-Torres, Alan Wong, Carole Escartin, and Julien
746 Valette. Characterizing extracellular diffusion properties using diffusion-weighted mrs of sucrose
747 injected in mouse brain. *NMR in Biomedicine*, 34(4):e4478, 2021. doi: [https://doi.org/10.1002/](https://doi.org/10.1002/nbm.4478)
748 [nbm.4478](https://doi.org/10.1002/nbm.4478). URL [https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/](https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.4478)
749 [10.1002/nbm.4478](https://doi.org/10.1002/nbm.4478).
- 750 Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Courn-
751 peau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der
752 Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson,
753 Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake
754 VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero,
755 Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mul-
756 bregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing
757 in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.
- 758 Hui Zhang, Torben Schneider, Claudia A. Wheeler-Kingshott, and Daniel C. Alexander. NODDI:
759 Practical in vivo neurite orientation dispersion and density imaging of the human brain.
760 *NeuroImage*, 61(4):1000 – 1016, 2012.
- 761 Karl Zilles. Brodmann: a pioneer of human brain mapping—his impact on concepts of cortical
762 organization. *Brain*, 141(11):3262–3278, October 2018. doi: 10.1093/brain/awy273. URL <https://doi.org/10.1093/brain/awy273>.