White Matter microstructural property decoding from gradient echo data using realistic white matter models

Renaud Hédouin*, Riccardo Metere*, Kwok-Shing Chan*, Christian Licht‡, Jeroen Mollink†, Anne-Marie van Cappellen van Walsum†, José P.Marques*

* Radboud University, Donders Institute for Brain, Cognition and Behaviour
6525 HR Nijmegen, Netherlands
† Radboudumc, Medical Imaging and Anatomy, Nijmegen, Netherlands
‡ Computer Assisted Clinical Medicine, Medical Faculty Mannheim, Heidelberg University, Germany

***Index Terms*—White matter models, Microstructural properties, Magnetic susceptibility, Deep learning network**

***Abstract*—The multi-echo gradient echo (ME-GRE) magnetic resonance signal evolution in white matter has a strong dependence on the orientation of myelinated axons in respect to the main static field. Although analytical solutions, based on the Hollow Cylinder Model have been able to predict some of the behaviour the hollow cylinder model, it has been shown that realistic models of white matter offer a better description of the signal behaviout observed.**

In this work, we present a pipeline to (i) generate realistic 2D white matter models with its microstructure based on real axon but with arbitrary fiber volume fraction (FVF) and g-ratio. We (ii) simulate their interaction with the static magnetic field to be able to simulate their MR signal. For the first time, we (iii) demonstrate that realistic 2D models can be used to simulate an MR signal that provides a good approximation of the signal obtained from a real 3D white matter model obtained using electron microscopy. We then (iv) demonstrate in silico that 2D WM models can be used to predict microstructural parameters in a robust way if multi-echo multi-orientation data is available and the main fiber orientation in each pixel is known using DTI. A Deep Learning Network was trained and characterized in its ability to recover the desired microstructural parameters such as FVF, g-ratio, free and bound water transverse relaxation and magnetic susceptibility. Finally, the network was trained to recover these micro-structural parameters from an ex-vivo dataset acquired in 9-orientations in respect to the magnetic field and 12 echo times. We demonstrate that this is an overdetermined problem and that as few as 3 orientations can already provide comparable results for some of the decoded metrics.

[Highlights] - A pipeline to generate realistic white matter models of arbitrary fiber volume fraction and g-ratio is presented; - We present a methodology to simulated the gradient echo signal from segmented 2D and 3D models of white matter, which takes into account the interaction of the static magnetic field with the anisotropic susceptibility of the myelin phospholipids; - Deep Learning Networks can be used to decode microstructural white matter parameters from the signal of multi-echo multi-orientation data;

I. Introduction

White matter (WM) consist mainly of myelinated axons and plays an important role for the transmission of information across the brain. The myelin sheath surrounding the axons acts as an electric insulator, thus increasing the transmission speed of the pulses. The development of myelin played a key role in evolution and the apparition of large vertebrate [1] and it is still central in brain maturation. The degradation of myelin, commonly referred to as demyelination, is present in various neurodegenerative diseases, and leads to severe motor and mental disabilities [2]. Such neurodegenerative disorders (e.g multiple sclerosis) show high variability among individuals, and it is difficult to predict and understand the course of the disease by only counting the number of lesions or comparing the values obtained in magnetic resonance (MR) relaxometry [3]. Therefore, non-invasive imaging methods that can investigate the WM microstructure and measure myelin properties may offer important means of studying neurodegenerative diseases, providing crucial information for diagnosis, monitoring progression and assessment of potential treatment effectiveness.

Direct MR imaging of the myelin is challenging due to the ultra-short transverse relaxation time of the phospholipid proton ($T_2^* = 0.3$ ms). Nevertheless, several attempts have been performed using zero or ultra short echo time techniques [4], [5]. Alternatively, myelin can be probed indirectly using magnetization transfer techniques [6], [7], multi-echo spin- or gradient- echo sequences [8]. However, the detection of myelin water remains challenging due to its short T_2 and T_2^* values ($\sim$10 ms). In this paper, we will focus on myelin water imaging using a multi-echo gradient echo (ME-GRE) sequence.

WM is a complex environment composed not only of axons but also different types of glial cells, vessels and more. However, the biophysical models typically used in magnetic resonance imaging (MRI) is simplified to 3 compartments: intra-axonal, myelin and extra-axonal water protons. Axons in WM have various shapes and sizes, with

a diameter ranging from 0.1 μm to 2 μm for unmyelinated axons and from 1 μm up to 10 μm for myelinated axons [9], but are typically modeled as cylinders. The myelin sheath, formed in the central nervous system (CNS) by oligodendrocytes represents approximately 80% of the brain's dry weight, consists of tightly packed phospholipid bi-layers united by the hydrophobic tails, separated by water layers [10]. These phospholipids because of their elongated form and their radial organisation around the axon have an anisotropic magnetic susceptibility [11], [12] which is diamagnetic when compared to surrounding water. These microstructural features are believed to be well approximated by a tensor with cylindrical symmetry. Thus, the susceptibility of myelin can be written as the sum of an isotropic (X_i) and anisotropic (X_a) component:

$$X = X_i + X_a = \chi_i \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \chi_a \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix} \tag{1}$$

where χ_i and χ_a are scalar isotropic and anisotropic susceptibility multiplicative constants, respectively. Various values have been reported in the literature of myelin for χ_i ranging from -0.13 to -0.06 ppm and χ_a ranging from -0.15 to -0.09 ppm [13]–[15] (with ppm considered with respect to the magnetic susceptibility of pure water).

In the presence of a strong magnetic field, a secondary microscopic magnetic field perturbation is created by these phospholipids [16]. This secondary field can be observed in both magnitude and phase of a multi-echo GRE signal [17]. One manifestation of the anisotropic magnetic susceptibility of myelin is that the MR signal of a GRE sequence shows a dependence on the orientation of the fibers relative to the main magnetic field. For example, it has been shown that simple T_2^* maps are orientation dependent [18], and hence unsuitable for the estimation of myelin properties. Part of this orientation dependence can be accounted for using *a priori* knowledge of fiber orientations [19]. More advanced biophysical modeling of the signal decay allows for the estimation of orientation dependent and independent components [20]. One alternative way to explore this complex interaction between magnetic field and the magnetic susceptibility of myelin, is myelin water imaging (MWI). In MWI the MR signal is fitted using a 3-pool model [21], [22], with each compartment having a specific relaxation time and frequency offset, yet this approach suffers from over-fitting issues [23], [24]. Using the hollow cylinder model (HCM) [25] and fiber orientation information derived from DWI, our group has shown that it is possible to overcome some of the ill-posedness nature of MWI using what was named Diffusion Informed Myelin Water imaging (DIMWI) [26]. Yet it has been previously shown that more complex and realistic WM models based on electron microscopy data [11] are better able to characterize the signal ME-GRE signal than the more simplistic HCM.

In this work, we present a novel approach to map some of the properties of WM microstructure by modeling the behavior of the MR signal of a ME-GRE imaging sequence measured for multiple orientation of the tissue sample and hence of the WM fiber bundles. For this, we first develop a method to generate a hypothetical 2D WM models based on realistic axon shapes. These models are then used to simulate the ME-GRE signal for different axon and myelin properties (notably, their relative size and volume fractions) and validated by comparing them to real 3D WM models based on electron microscopy data. These simulations are used to construct a dictionary of the ME-GRE, covering a wide range of WM properties. Finally the dictionary was used to train a deep neural network to map ME-GRE acquired signal using multiple orientations of a sample in respect to the magnetic field into white matter properties. We tested this deep learning method in various scenarios both *in silico* and *in vitro*.

II. Methods

A. 2D WM model

The magnetic susceptibility of myelin relative to its surrounding creates a magnetic field, that although small, affects the MRI signal both in phase and magnitude. This phenomenon have been used in the past to study WM orientation [19], [25] and can be studied both analytically and numerically considering various simplified WM models.

1) Hollow cylinder model: The HCM, proposed by Wharton and Bowtell, is commonly used to approximate WM microstructure [14]. The myelin sheath is represented by an infinite hollow cylinder with an inner radius r_i and an outer radius r_o. The inner part of the hollow cylinder is the intra-axonal compartment and the external part is referred as the extra-axonal compartment.

This cylindrical representation of WM into 3 compartments allows an analytical derivation of the field perturbation in each of those regions and characterization WM using:

- Fiber volume fraction (FVF) - the proportion of myelinated axon within the model
- g-ratio - the ratio between the intra-axonal radius (r_i) and the myelinated axon radius (r_0):

$$\text{g-ratio} = \frac{r_i}{r_o} \tag{2}$$

This solution, very convenient to model, offers for example an analytical estimation of the fiber-orientation dependence of $R2^*(1/T2^*)$ map [25].

However, it has been recently demonstrated that the HCM has intrinsic biases compared to a more realistic WM model created from electron microscopy data [11]. The circular axon shapes create artificially large frequency peaks, in particular within the intra-axonal compartment, which are not present in a realistic model. In the following section we will present the creation of a realistic 2D WM model based on real axon shapes and realistic size distributions.

2) Electron microscopy based models: In this study, we used an 2D electron microscopy image of an entire slice of a canine spinal cord from an histology open database [1] as our database of axon shapes. The sample is 5mm width and 7.5mm long with a 0.25μm resolution which corresponds to a 20.000×30.000 image. An open-source segmentation software was used to segment the image that lead to a collection of ~ 600.000 myelinated axon shapes [27]. The resolution is sufficient because we do not want to segment unmyelinated axons that have been shown to have no significant impact in the obtained ME-GRE signal [11]. The unmyelinated axons are therefore included within the extra-axonal space. In case of a realistic axon shape, the g-ratio is redefined as the square root of the ratio between the intra-axonal surface and the outer surface (measured as the number of myelinated pixels with at least one side in direct contact with intra or extra-axonal space).

3) Axon packing algorithm: A set 400 of axon shapes were randomly picked from the collection above to create a realistic 2D WM model with predefined FVF and g-ratio. To do so, we developed an axon packing algorithm based on an existing software [28] that had been initially developed for cylindrical axon models. The packing process is performed as follow (see Fig 1):

Algorithm 1: Axon packing

Data: Set of N myelinated axon shapes
Initialization: N axons equally spaced on a grid
current FVF = initial FVF
while *current FVF < maximum FVF* **do**
  Axons are attracted to the grid center
  Axons which overlap repulse each other
  current FVF = FVF within a mask
end

In the current implementation, as the axon shapes are picked randomly, they do not necessarily fit optimally together (during the attraction and repulsion process, the axon is not allowed to rotate) which creates small gaps within the model. The maximum FVF parameter, corresponding to a model where the axons are highly packed while avoiding overlap was empirically found to be 0.85. According to literature, such an FVF value already represents a WM model with a very high axon density [29].

4) Obtaining an expected FVF: Once the maximum FVF for a given collection of axons is achieved, this packed WM model was used to obtain a new model with an a different FVF. Two different methods, illustrated Fig 1, were proposed: (i) randomly remove axons or (ii) spread the axons from the figure center. The first method creates important gaps within the extra-axonal space that could correspond to glial cells or bundles of unmyelinated axons, while the second method creates a more uniformly distributed WM model. Based on the EM data visually explored up to now, both could be valid representations. Their corresponding field perturbation histograms were close enough and both models were used to enforce the diversity of our WM model dictionaries.

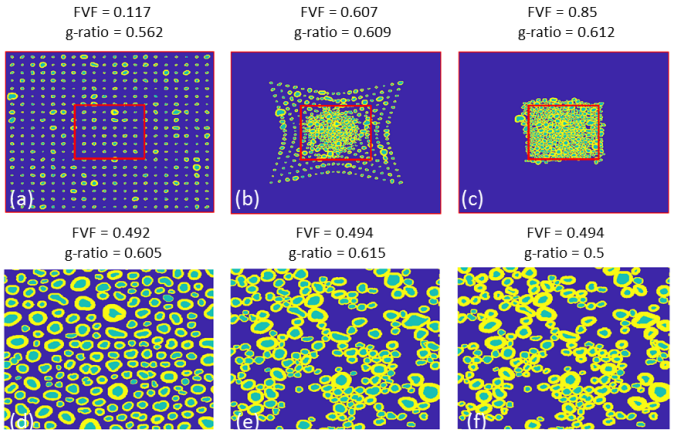


Fig. 1. Top row: 400 axons are placed on a grid (a) and packed following an attraction/repulsion method (b) until to reach high FVF (c). Bottom row: Zoom on the mask delineated by the red square. A desired FVF is reached spreading the axons from the center (d) or randomly removing some axons (e). Keeping the same axons and thus the same FVF, the myelin thickness can be modified to obtain an expected g-ratio (f)

5) Change the g-ratio: Finally, the mean g-ratio of the model was modified, while keeping the FVF constant. This operation was performed on an axon-by-axon basis by dilating or eroding the inner myelin sheath by one pixel depending on whether the g-ratio was to be decreased or increased. Each axon has a given probability to be randomly picked, this probability is linked to its diameter. As the dilatation/erosion is fixed to one pixel, larger axons need to be picked more frequently to respect the original proportion of FVF. The modification of the g-ratio is illustrated in Fig 1. Eventually, different models with similar FVF and g-ratio can be created using our large axon shapes database and the code made available in the toolbox.

B. Signal creation

To be able to use these 2D models to simulate the ME-GRE signal, we need to define the susceptibility of pixel element, compute the induced magnetic field perturbation and eventually simulate the signal evolution in this inhomogeneous environment.

1) Magnetic susceptibilities: The susceptibility tensor X_R within the myelin sheath in a 2D model is determined by the phospholipid orientations ϕ on that plane:

$$X_R = R_z(\phi) \cdot X \cdot R_z(\phi) = X_i + R_z(\phi) \cdot X_a \cdot R_z(\phi) \quad (3)$$

with $R_z(\phi)$ the 3D rotation matrix around the z axis. In simple cases, as for the HCM, the computation of ϕ is trivial. However, for more complex axon shapes, there is no proper definition of a radial orientation between two boundaries. The orientation of the phospholipids is

[1] https://osf.io/sgbm8/

estimated on an axon-by-axon basis. First, the selected axon is placed in a small matrix (including 10 pixels of each side of the axon edges for computational time considerations), then the extra-axonal, myelin and intra-axonal compartments are given the values of 0,1 and 2 respectively. The resulting map is smoothed with a 2D Gaussian filter with a width of 5×5 to create a smoothed pyramidal structure, if the myelin sheath is too large and still contains piecewise constant part after smoothing, the process is repeated, and finally a gradient direction map is computed. As the map is smoothly varying from 0 to 2 within the myelin compartment, the gradient at each point will define the steepest direction from the extra- to the intra-axonal space and should follow the phospholipid orientation (see Fig 2).

2) Field perturbation: From the phospholipid orientation map, the susceptibility tensor map can be calculated using Eq. 3. The susceptibility tensor map is used to compute the field perturbation in the frequency domain as described in [30]. An illustration of the field perturbation generated by a single axon with several B_0 orientations is shown in Fig 2. The induced field perturbation strongly depends on the B_0 orientation. A magnetic field parallel to the axon orientation only creates a small negative field shift within the myelin sheath while a perpendicular magnetic field creates much stronger perturbations within the 3 compartments. The overlapping frequency spectra of the 3 compartments make them hard to disentangle.

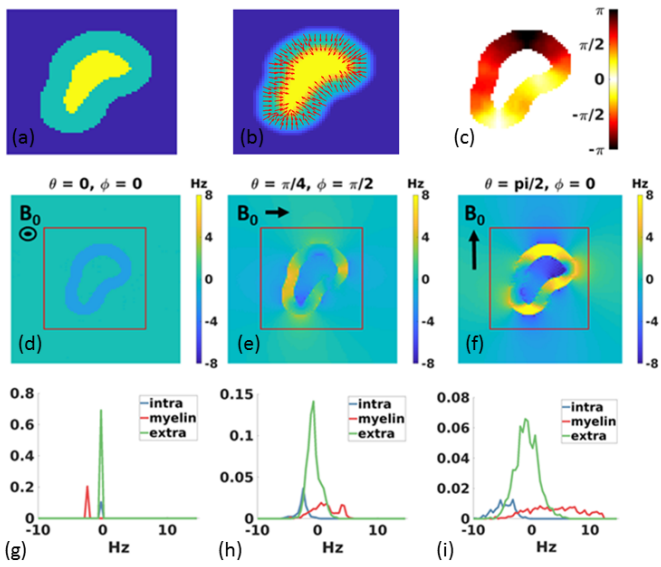


Fig. 2. First row: Phospholipid orientation estimation. (a): Original axon, the extra-axonal part is filled with 0, the myelin with 1 and the intra-axonal with 2. (b): Model smoothed with a Gaussian filter. Right : (c) Gradient orientation computed on the smoothed map. Second row: Field perturbation for one axon with 3 different magnetic field orientations. Third row: Corresponding histograms computed within the red square to keep a reasonable FVF.

3) ME-GRE signal: ME-GRE signals are computed as:

TABLE I
MIDDLE COLUMN: PARAMETER RANGE USED IN OUR DICTIONARY. THIRD COLUMN: USUAL PARAMETER VALUES FOUND IN WM. *THE RELATIVE WATER WEIGHT DEPEND ON THE ACQUISITION PARAMETERS, FLIP ANGLE AND TR, SUBSEQUENTLY THERE IS NO USUAL VALUE AND THE ONE PRESENTED IS USED WITHIN THE TYPICAL WM DEEP LEARNING EXPERIMENT.

Model parameters	Dictionary	Typical WM values
FVF	0.1:0.1:0.8	0.7 [a]
g-ratio	0.5:0.05:0.85	0.65 [b]
χ_i	-0.2:0.1:0.2	-0.1 [c]
χ_a	-0.1 (fixed)	-0.1 [a]
$T_{2,Intra-Extra}$	20:20:100	60 [d]
$T_{2,Myelin}$	4:4:20	16 [d]
$w_{Myelin}/w_{IntraExtra}$	0.5:0.5:3	2*
Fiber orientations	20	/

[a] (Choy et al., 2020) [31]
[b] (Mohammadi et al., 2015) [32]
[c] (Wharton et al., 2012) [14]
[d] (Xu et al., 2018) [11]

$$S(t) = \sum_{n=1}^{3} \left(w_n \exp\left(\frac{-t}{T_{2,n}}\right) \sum_r \exp\left(-it\gamma\Delta B_n(r)\right) \right) \quad (4)$$

where each compartment has a specific transverse relaxation $T_{2,n}$, a water weight w_n reflecting the water signal, which includes proton density and T_1 saturation effects, and a corresponding field perturbation $\Delta B_n(r)$. An illustration of the ME-GRE signals simulated using Eq 4 is shown with two examples of WM geometry in Fig 3.

MRI data amplitude depends, not only on the magnetization amplitude, but also on the RF coil sensitivity and receiver gain. The phase depends on the RF transceiver and on the quality of the B0 shimming and presence of fields due to the susceptibility of neighbouring pixels. To be able to compare our signal simulations to real data, this signal (and the real data) is normalized as follow:

$$|\hat{S}(t)| = |S(t)|/|S(1)| \quad (5)$$

$$\arg(\hat{S}(t)) = \arg(S(t)) - p_1 - p_2 \times t \quad (6)$$

where $\arg(\hat{S}(t))$ is the phase of the signal and p_1 and p_2 are the coefficients of $p(t) = p_1 + p_2 \times t$, the line that best fits the original phase $\arg(S(t))$. This normalization is done to keep the second order evolution while removing the linear part that corresponds to the phase and frequency offset of the MRI acquisition which are hardware-dependent.

4) Model validation: While the realistic 2D WM models have been shown to better represent the ME-GRE signal of WM than the simple HCM, they assume the replication of the same structure along the third dimension resulting in bundles that are unrealistically aligned and cannot represent the natural dispersion present in a real axon bundle. Dispersion can occur not only in regions of fiber crossing,

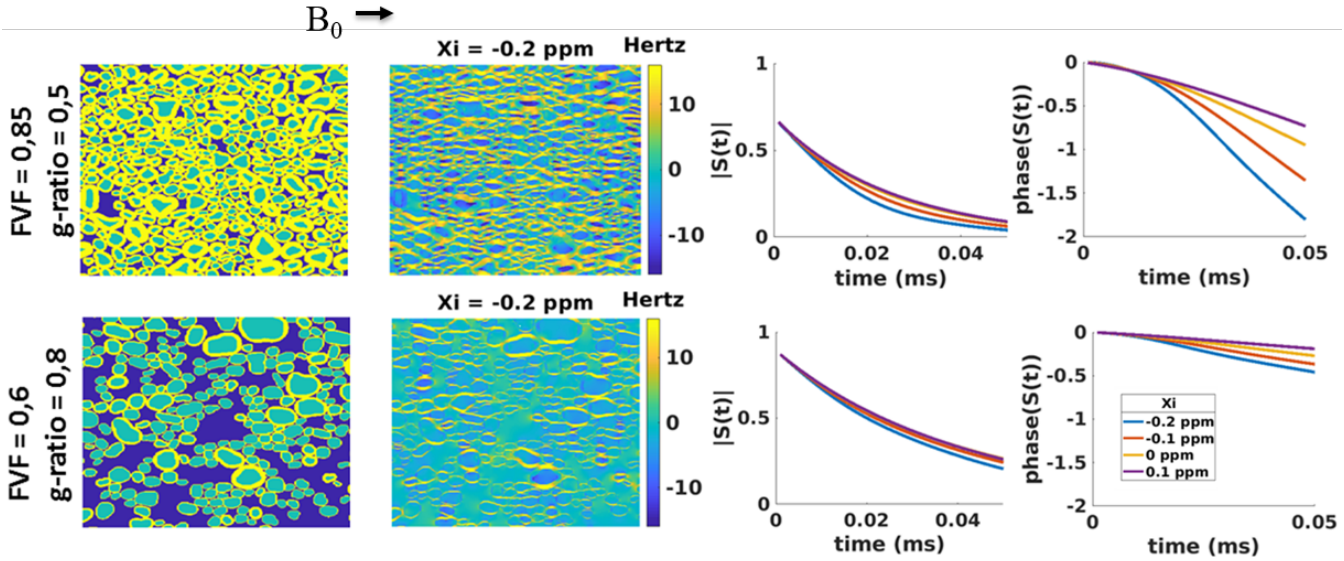


Fig. 3. WM models with different FVF and g-ratio and their corresponding field perturbations. The ME-GRE signal magnitude and phase are presented for 4 difference χ_i values. The other parameters are fixed according to literature values (see table I.

fiber kissing, but also in regions traditionally expected to be unidirectional such as the corpus callosum [33]. However, 3D models are hard to construct, not only because of the lack of 3D EM data (that could represent a ground truth), but also because of the complexity of 3D axon packing [34]. Also, it could be the case that the 2D axon shapes, used in our realistic WM modeling, are elongated due to being obtained from cutting through axons that were not perpendicular to the surface. Furthermore, in the case of our application, the estimation of the susceptibility tensor map and the field perturbation in 3D models would make the process even more time consuming. We have designed a small study, presented in the Appendix , to evaluate the ability of our 2D models to represent a real 3D model with comparable microstructural properties.

C. Dictionary creation

A dictionary of signal evolution can be created using the simulated ME-GRE signals in the presence of different WM model. Such dictionary can be used to derive the microstructural tissue properties from the ME-GRE signal by root mean square minimization, as done for example in fingerprinting [35]. Alternatively, a deep learning network can be trained to learn the tissue properties from the dictionary as will be demonstrated later.

The WM model and the magnetic field distributions present on each of its compartments depends on 5 microstructure related parameters: FVF, g-ratio, χ_i and χ_a, as well as the fiber orientation. For the purpose of training a deep learning network, we considered repeating simulations with various axon packing using the same properties as aforementioned. The ME-GRE signal from each WM model depends on the specific NMR properties of each compartment $(w_n, T_{2,n})$. This would result in 6 supplementary parameters. To minimize dictionary size, the T2s and water weights of the intra-axonal and extra-axonal compartments were defined equals and the signal amplitudes was always normalized (so that $\sum w_n = 1$), reducing the number of parameters from 6 to 3: $w = S_{0,IntraExtra}/S_{0,Myelin}$ referred as the relative water weight, $T_{2,Myelin}$, $T_{2,Intra-Extra}$. The parameter ranges, used to construct the dictionary, are presented in Table I along with typical WM values. The dictionary has 8 dimensions, with 5 to 20 entries per dimension leading to 7.680.000 vectors. In the following *in silico* and *ex vivo* experiments, all the dictionaries have those same parameter ranges.

Each entry of the dictionary is composed by the normalized signal magnitude and phase (or real and imaginary components, 2 x nTE with nTE the number of echo times in the simulation) and an additional entry encoding the fiber orientation information characterized by the angle between the fiber and the static magnetic field. When deriving the microstructural properties from measurements with multiple orientations with respect to the magnetic field, the signal is concatenated along the n orientations which leads to a vector size of $n \cdot (2\text{TE}+1)$. An illustration of such simulated normalized signals magnitude and phase with different orientations is presented Fig 4. Conversely to a single orientation dictionary, this multi-orientation dictionary is only valid for a specific set of rotations used in a specific acquisition.

D. Deep Learning

The ME-GRE signal dictionary was used to train a deep learning network using Keras [36]. For all the following experiments, the dictionaries were trained on 7 entire sets of WM models and evaluated by the loss function on another set of WM models, which correspond to a validation split of 0.125. This network is composed with 3

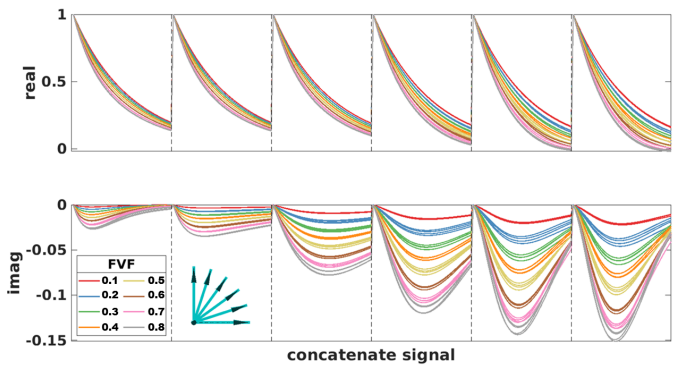


Fig. 4. The ME-GRE signal is simulated with 6 magnetic field orientations, $\theta = 0 - \pi/2$ equally space (see arrows), for WM models with different FVF from 0.1 to 0.8 (4 models for each FVF). The top and bottom row represents respectively the signal real and imaginary part for each 6 orientations separated by a vertical black line.

hidden layers of size $2*l_i*l_o$, $1.5*l_i*l_o$, $1.25*l_i*l_o$, l_i and l_o being the concatenate signal length and the number of parameters, with a respective drop out of 0.4, 0.2, 0.1 using an tanh activation function and an additional linear layer. Both inputs and outputs were normalized, a stochastic gradient descent optimizer was used and the loss function is a mean absolute error.

To gain experience on the ability and limitations of our network to derive microstructure properties, its performance was first tested on simulated data. Particularly we wanted to evaluate what was the optimum echo time range and the number of echoes, as well as study the gains associated with different numbers of sample rotations needed to successfully recover WM properties (which will affect our data acquisition protocol). The design and training of the network was also subject of careful attention. The deep learning hyperparameters were tuned following an empirical approach, with the chosen ones giving both accurate results and robust to the change of signal parameters.

The validation loss function (mean absolute error of the parameters estimated on a validation data set - one set of WM models which is not used for training) was used as a metric to evaluate the convergence of the network. All the parameters, within their range, were re-scaled between 0 and 1, to make validation loss a less arbitrary number. This metric is an average of the mean absolute error for each parameter, thus, it does not allow to make fine comparisons. Despite this remark, the validation loss is a classic and robust way to evaluate the training process with an unique number.

1) Deep Learning performance evaluation on simulated data: The robustness of the parameter recovery was tested by adding a complex white noise (0 %, 0.5 %, 1 %, 2 % and 4 %) to a ME-GRE signal on a dictionary used in the training and validation processes. The first 3 columns of Table II summarize the parameters used in the creation of the dictionary and training of the network. The rotations used were chosen to mimic the experimental protocol used on an *ex vivo* acquisition described later in this section. The noise levels mentioned above are relative to the signal amplitude at the first echo, TE=2.15 ms

The ME-GRE signal of a given white matter model depends on the magnetic field orientation in respect to its structure (see Fig 4), this lead us to adopt a multi-orientations approach when trying to decode WM microstructure properties. However, as an increased number of orientations means a longer acquisition time, we performed a theoretical comparison study to estimate the benefit of using a large number of orientations vs a reduced number of orientations with higher SNR. A dictionary with 16 optimal orientations was created for 3 different noise levels (0,1 and 2%). In order to maximize information, each fiber should have the largest possible range of θ from 0 to $\pi/2$. To do so, the 16 3D rotations had evenly spread axis on the sphere with a common $\pi/2$ angle. Then, for a range of number of orientations from 1 to 16, a subset of this dictionary was used to train a deep learning network.

The influence of the number of echoes on the deep learning parameter recovery performance was tested. To do so, several networks were trained with a fixed echo spacing (3.05 ms - mimicking our experimental protocol), a various number of TE (5, 10, 15, 20, 25 and 30) and noise levels. At this stage no considerations of the impact on T_1 weighting were included on the analysis.

Finally, we tested the deep learning for one set of realistic parameter values of WM (see Table II), that allows to detail the behavior of each parameter individually. The signal was simulated 125 times for 8 independent white matter models leading to 1000 signal simulations with each different noise level. We tested two methods to recover the parameters: (i) using a deep learning network trained with a noise matching to the simulated noise; using a deep learning trained with a maximum noise level regardless of the simulated signal noise.

E. Ex vivo data acquisition

A formalin fixed post-mortem brain (female, 88 years old, 26 hours of post-mortem interval and 7-month fixation period) was scanned in a 3T scanner (Siemens, Prismafit). The brain was scanned in 9 orientations relative to the static magnetic field. To avoid brain deformation between different rotations, a customised 3D brain holder was built and used throughout the scanning session, see Fig. 5. Prior to scanning, formalin was washed out in distilled water and prepared in low pressure environment, using a vacuum pump at 20mBar during 12h to remove all air bubbles trapped in the various cortical sulci. During this period the brain was occasionally rotated to ensure removal of air trapped inside the ventricles.

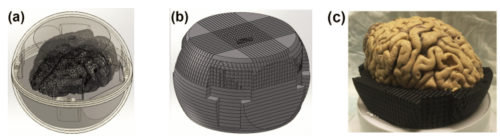


Fig. 5. Brain holder: (a) in its sphere, (b) alone, (c) open

TABLE II
THIS TABLE DESCRIBES THE DICTIONARY PARAMETERS, (TEs, ROTATIONS, NOISE LEVEL, NUMBER OF MODELS) AND A DEEP LEARNING PARAMETER (NUMBER OF EPOCHS) ASSOCIATED WITH EACH EXPERIMENT. FOUR FIRST COLUMNS: *in silico* EXPERIMENT, LAST COLUMN: *ex vivo* EXPERIMENT.

Parameter \ Experiment	Epochs dependence	TEs dependence	Rotation dependence	Typical WM	*Ex vivo* data
TEs	2.15-3.05-35.7	1.8-3.2-14.6/94.6	2.4-4.4-50.7	2.15-3.05-35.7	2.15-3.05-35.7
Rotations	9	6	1 to 16	9	9
Noise level	$0, 0.5, 1, 2, 4\%$	$0, 1, 2\%$	$0, 1, 2\%$	$0, 0.5, 1, 2, 4\%$	4%
Number of models	8	8	8	8	8
Epochs	40	20	40	40	40

For each head positions the following protocol was repeated: - (a) 3D monopolar ME-GRE with 12 echos (TE = 1.7 : 3.05 : 35.25ms, TR = 38 ms), with a 1.8mm isotropic resolution and Matrix size (128x128x128), acquisition time 8.21 mins. This protocol was repeated 6 times with flip angles $\alpha = 5°/10°/15°/20°/35°/65°$); - (b) an MP2RAGE with 1mm isotropic resolution was acquired for co-registration purposes. The MP2RAGE parameters were adapted to be able to map the short T_1 values present in fixed tissue (TR/TI1/TI2=3s/0.311s/1.6s α_1/α_2); Finally, for the last sample position, DWI protocol was acquired to provide fiber orientation information (TR/TE = 3.78s/71.2ms, 256 diffusion-encoding gradient directions, b = $2500s/mm^2$). Because the formalin fixation process and the reduced temperature of the sample compared to *in vivo* (Room Temperature $\simeq$ 23°) significantly reduce water diffusivity, the protocol was repeated 20 times to achieve robust fiber orientation information.

F. Ex vivo data processing

Each of the 9 MP2RAGE images from the 9 brain rotations were co-registered to a reference position using FLIRT from fsl [37]. Corresponding transformations were then applied to the ME-GRE data (magnitude and unwrapped phase separately), finally the registered data were normalized following Eq 6. A DTI was estimated for each DWI and the 20 DTIs were averaged using a log-Euclidean framework [38]. Eventually, the fiber orientation was defined as the main orientation of the average tensor.

A ME-GRE dictionary was simulated for this particular acquisition, and the corresponding deep learning network was trained using the parameter ranges described in Table I and II. Finally, the microstructure parameter maps (FVF, g-ratio, χ_i, $T_2, Myelin$. $T_2, Intra - Extra$, and the relative water weight) were estimated individually for each set of flip angles. This resulted in 6 independent sets parameter maps, where only the relative water weight term is expected to vary across acquisitions. It was thus possible to compute the mean and standard deviation of the microstructure parameter maps that were expected to remain constant across flip angles to estimate the precision of those measurements.

Finally, the last experiment was performed by using a restricted number of rotations that can be achived during an *in vivo* experiment. Among the 84 possible combinations of 3 rotations chosen within the original 9 rotations, the 10 ones that insured the largest fiber orientations ranges were picked. The subsets of *ex vivo* data for the 10 combinations of 3 rotations with a flip angle of 20°, the corresponding dictionaries, and deep learning networks were created, leading to 10 entire sets of brain parameter maps. This was used to compute the mean and standard deviation across different combinations of 3 rotations. Finally, the absolute difference maps between the mean parameter maps with 3 rotations and the original ones with 9 rotations, both with a flip angle of 20°, were estimated.

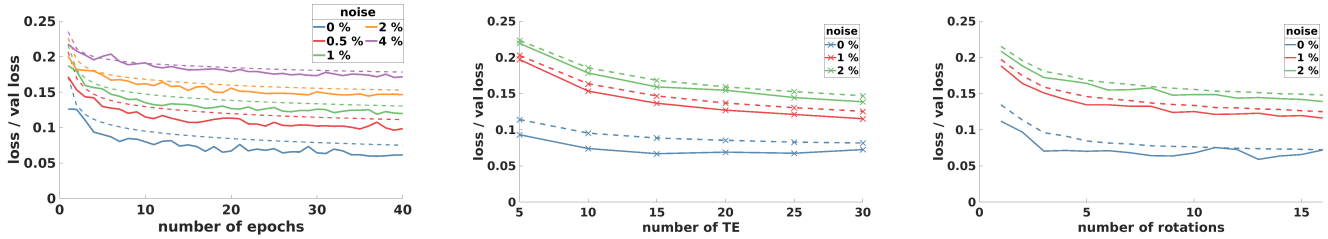
III. RESULTS

A. Deep learning performance on simulated data

1) Noise level: Figure 6(a) shows the dependence of the loss function of the deep learning network for 5 different noise levels as a function of the number of epochs used. After a fast drop during the first 3-5 epochs, the loss function continues a slow decay, reaching a plateau for the noisier signals. Interestingly, the loss functions on the test data (solid lines) are slightly lower than on validation data (dashed line). This difference is attributed to the fact that the validation loss function is averaged along the entire epoch, whilst the test loss function is computed at the end of each epoch. From this analysis we concluded that 20 epochs should be a good compromised between training efficiency and parameter recovery.

2) Echo times: Fig 6(b) presents the dependence of the loss function on the number of echo times used. It shows that the wider the range of echo times the lower the loss function is. The loss function clearly improves between 5 to 15 TE (49ms), but its improvement is smaller after that, even if a plateau has never been totally reached for signal with noise even after 30 echos. Our simulations did not include any echo time dependent noise, arising from physiological noise or scanner drifts, which are common in gradient echo acquisitions, and would make later echo times less useful for decoding. We postulated that 20 echos would be sufficient for an experimental protocol.

3) Number of magnetic field orientations: Fig 6(c) shows that, as expected, the loss functions decreases when increasing the number of rotations for all noise levels. Note that in interest of computation time, the subset of rotations might not be optimal for all number of rotations



(a) (b) (c)

Fig. 6. Deep learning training evolution for different noise levels relative to several acquisition parameters. The solid line is the loss function whilst the dashed line is the validation loss function, that represents the same mean absolute error respectively computed on the train and on the test data set. (a) Training along the number of epochs. (b) Training along the number of echoes. (c): Training along the number of rotations.

tested (as a subset of the initial 16 orientations was used). Furthermore the specific number/set of rotations depends on the orientation of the fiber of interest. The deep learning benefits from the first 3-6 distinct rotations, similar to what has been demonstrated for Susceptibility tensor imaging [39] and for fiber orientation mapping [14], and plateaus after this. In a given acquisition time we can either decide to have an improved SNR per orientation or increased number of rotations. When moving from 1 to 2%SNR levels this corresponds to an decrease of the acquisition time by a factor 4 or number of rotations. Thus 16 orientations at 2% noise could be acquired in the same time as 4 orientations at 1% noise level. It can therefore be concluded that there is a benefit in maximizing the number of orientations beyond 5 as the loss function for 16 rotations at 2% noise was the same as that of 6 orientations and 1% noise. In our acquisitions, we used 9 to 10 orientations, to avoid excessive acceleration of each orientation, as this could bring parallel imaging artifacts into play when trying to further reduce the acquisition per orientation.

4) Selective set of parameters: Fig 7 shows the performance of the deep learning networks to recover the various microstructural parameters of what could be considered a typical white matter model. Although the average recovered parameters are closed to the original ones regardless the signal noise level many of the differences would be statistically significant. Particularly, the relative water weight suffers a constant positive bias for all networks and simulated signal. Surprisingly, the standard deviation for all parameters (excluding χ and $T_{2,Myelin}$) is considerably lower when the deep learning was trained with a 4% noise level rather then the matched noise level. Thus, a dictionary with a high noise level was used in our *ex vivo* experiment presented in the following. When comparing the width of the various distributions, compared to the range used in the training the network (see Table I), the valued of χ, g-ratio and relative water weight are likely to have the largest biases and noise.

B. Ex vivo experiment

The 6 microstructure parameter maps obtained from the *ex vivo* brain with a flip angle of 35 are presented in Fig 8, along reference images for coronal and sagittal views from the ME-GRE and MP2RAGE for visual comparison with more standard contrasts. The microstructure parameters were computed with a network with a 4% noise level. White matter is clearly discernible from grey matter and deep gray matter on the FVF and relative water weight maps. It should also be noted that FVF and the intra and extra axonal T2 have a very strong contrast between white matter and deep gray matter (although the latter has reduced contrast between grey and white matter). That observation is particularly interesting because it suggests that with our modeling we were able to remove myelin contributions to the T_2 contrast. On the other hand, the g-ratio map and T_2 of myelin seem to have large contrast within white matter which are expected to vary within the brain. The sagittal maps show that an higher FVF, lower T_2 of myelin and lower g-ratio in the corpus callosum compared to the rest of the brain. Interestingly, CSF presents an almost null FVF along with a high T_2 intra/extra axonal, which is to be expected as there are no structures generating an anisotropic signal evolution in this region. The χ_i values of myelin within WM are slightly positive where a negative value is expected, this effect could be due to the fixation process.

In addition to considering one particular flip angle, as in Fig. 8 , Figure 9 shows the mean and standard deviation across the 6 flip angles used. Most microstructure parameters should not depend on the flip angle, except the relative water weight that is related to the proton density as well as the TR and the flip angle. The mean parameter maps have globally the same characteristics than the ones presented previously with a flip angle 35, see Fig 9. The corresponding standard deviation maps reveal low values, showing a good robustness of the parameter recovery for several acquisitions. Finally, the relative water weight maps are shown in the same figure. As expected, an increase along the 6 flip angles is visible, in particular for the 2 higher flip angles, 35 and 60. Reflecting the fact that the T_1 of myelin water is significantly shorter than that of free water.

Fig 10 shows the brain parameter maps computed from several data subsets each using a different combinations of 3 rotations. The mean parameter maps highlight the expected brain structures, such as CSF, deep gray matter, WM, GM. Yet, the contrast seems lower compare to the

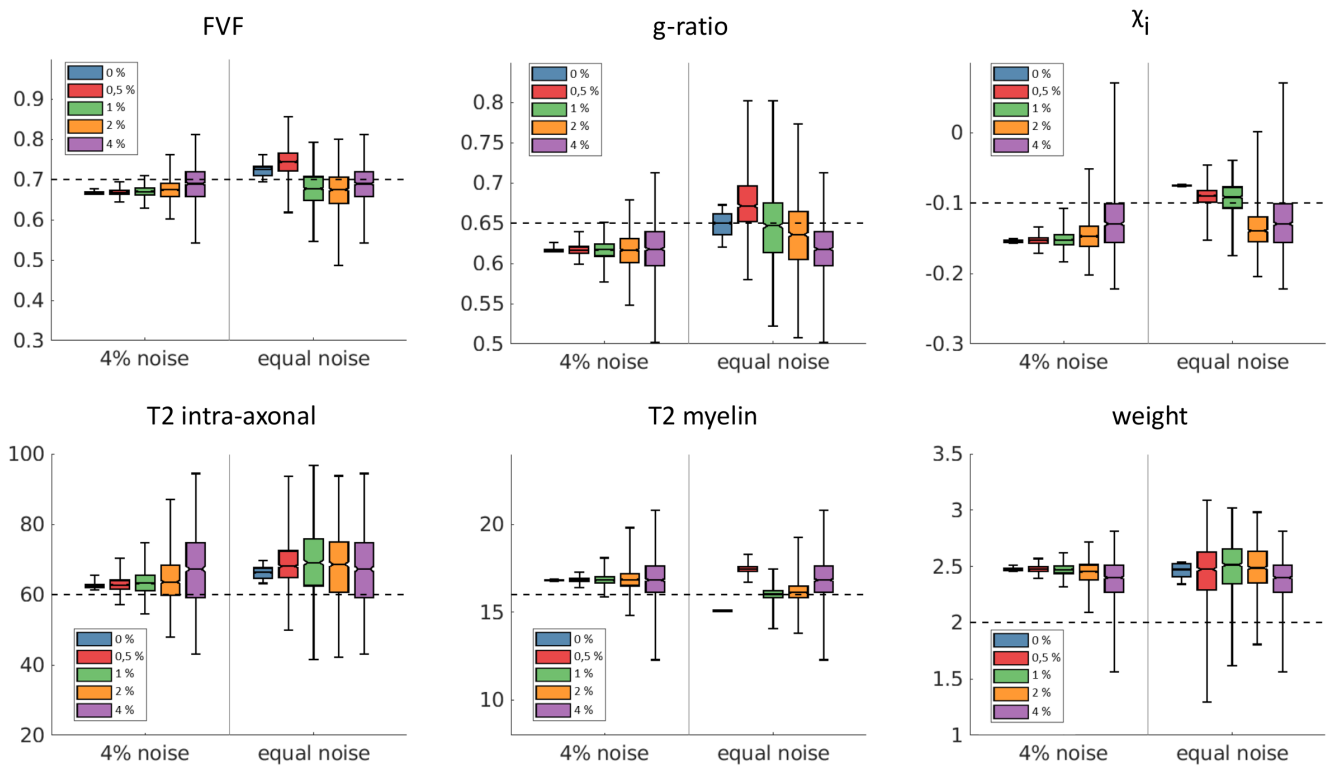


Fig. 7. Each box represents the estimation of one parameter recovery for 5 different signal noise levels $(0\%, 0.5\%, 1\%, 2\%, 4\%)$. Within a box, the left side use a single deep learning trained with 4% noise regardless of the noise level while the right side use 5 deep learning, each one trained with a noise equal to the signal level.

parameter maps obtained with 9 rotations, in particular within deep gray matter, as illustrated by the absolute difference maps. The standard deviation maps, estimated across 10 combinations of 3 rotations, reveal very low values. Thus, the process seems very robust when the same number of rotations is considered.

IV. Discussion

A. *White matter models: promise and limitations*

We introduced a pipeline to create a simple but realistic biophysical model to simulate the MRI ME-GRE signal. These WM models contain real axonal shapes and a g-ratio variability similar to what is reported in tissue samples (data not shown), and have varying levels of fiber volume fraction within themselves as a result from the axon packing approach. Yet, some effects are explicitly overlooked: (1) diffusion within the compartments, (2) chemical exchange and (3) other sources of susceptibility perturbations beyond the myelin sheath. Diffusion has been demonstrated to have a minor effect for white matter models based on EM data [11] when compared to the HCM or simple cylindrical perturbers [40]. Chemical exchange between myelin water and myelin protons results in frequency shift, and thus, can be accounted for by adding an exchange term in the HCM [25]. A recent work based on Generalized Lorentzian Tensor Approach directly considers the water layer within the myelin to probe this exchange [41]. The size of this frequency offset term has been reported to be of 0.02 ppm in the corpus callosum [14], but models have been proposed that would make this offset depend on the number of myelin layers and therefore vary throughout the brain and fibre bundles [42]. Yet, chemical exchange has been demonstrated to have a larger impact when measuring the longitudinal relaxation in white matter, which is an aspect that, for the sake of complexity, we have not included in our dictionaries. The extra-axonal compartment currently includes everything that is found outside of the axon. More classes with specific properties could be used, particularly: free water (CSF and interstitial spaces); blood vessels; bound-water compartment (that represents the water bound to macromolecules present in cell walls and organelles [43]), and iron accumulated in ferritin, amongst other. Blood vessels occupy a very small fraction of tissue volume (1-4% in WM and GM, but deoxygenated (venous) blood has a much larger susceptibility difference to free water than myelin) and tend to follow the orientation of white matter axon bundles [18]. This is expected to introduce some degree of T_2^* anisotropy that would act as a confound in our *ex vivo* experiment. Ferritin, which is known to be strongly paramagnetic, can be found everywhere in the brain (with increasing quantities found from WM, GM to deep gray matter where it can be found in large quantities [44]). On our current implementation, iron is expected to be equally distributed in the intra and extra-axonal space. As a result ferritin will be mapped as a reduction of the intra and

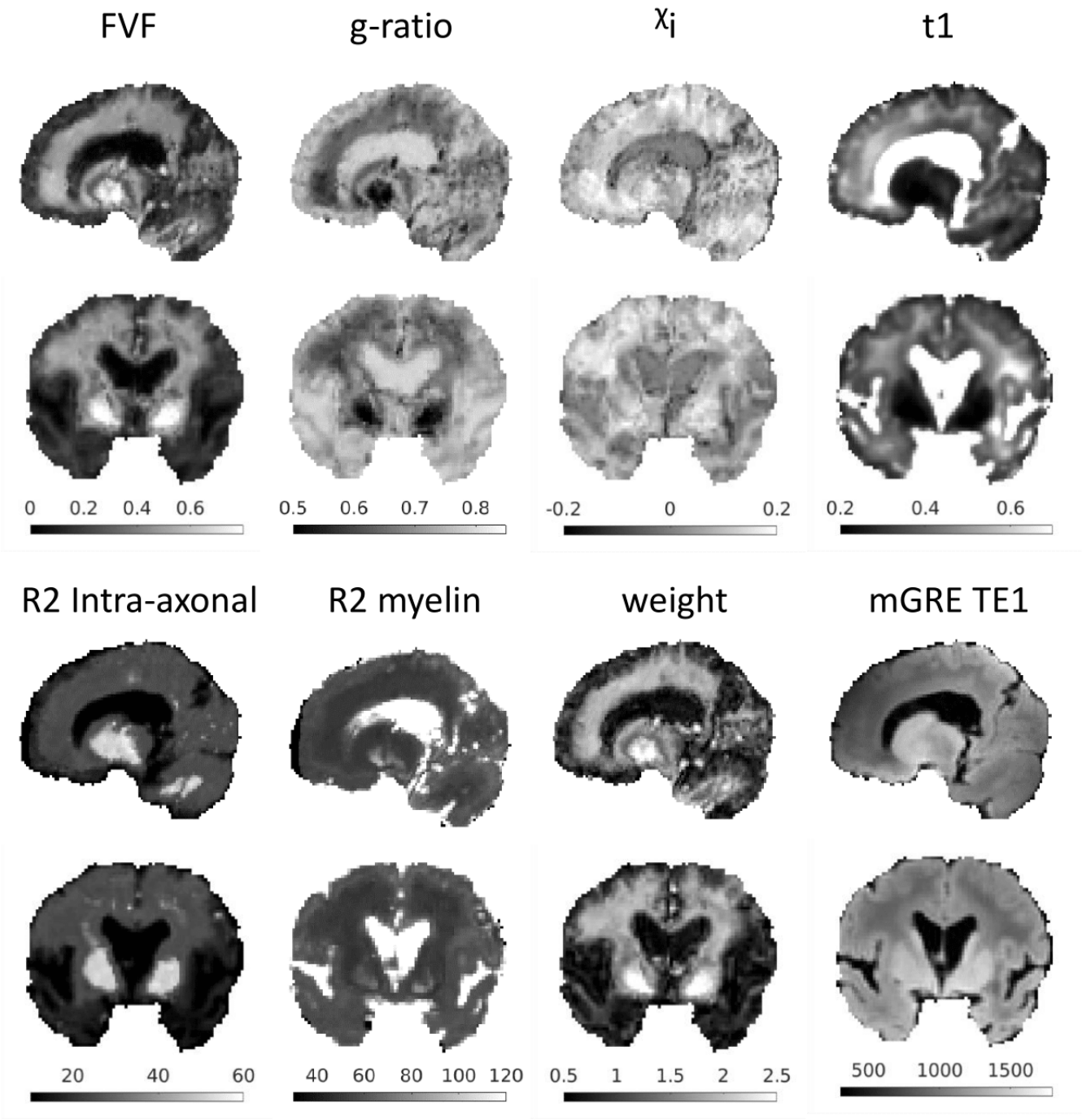


Fig. 8. Brain parameter maps estimated from *ex vivo* acquisition with flip angle 35 in a sagittal slice cutting the corpus callosum and transverse to coronal slice cutting through the globus pallius. Additional T_1 maps, estimated from MP2RAGE, and ME-GRE magnitude first echo provide structural information for comparison.

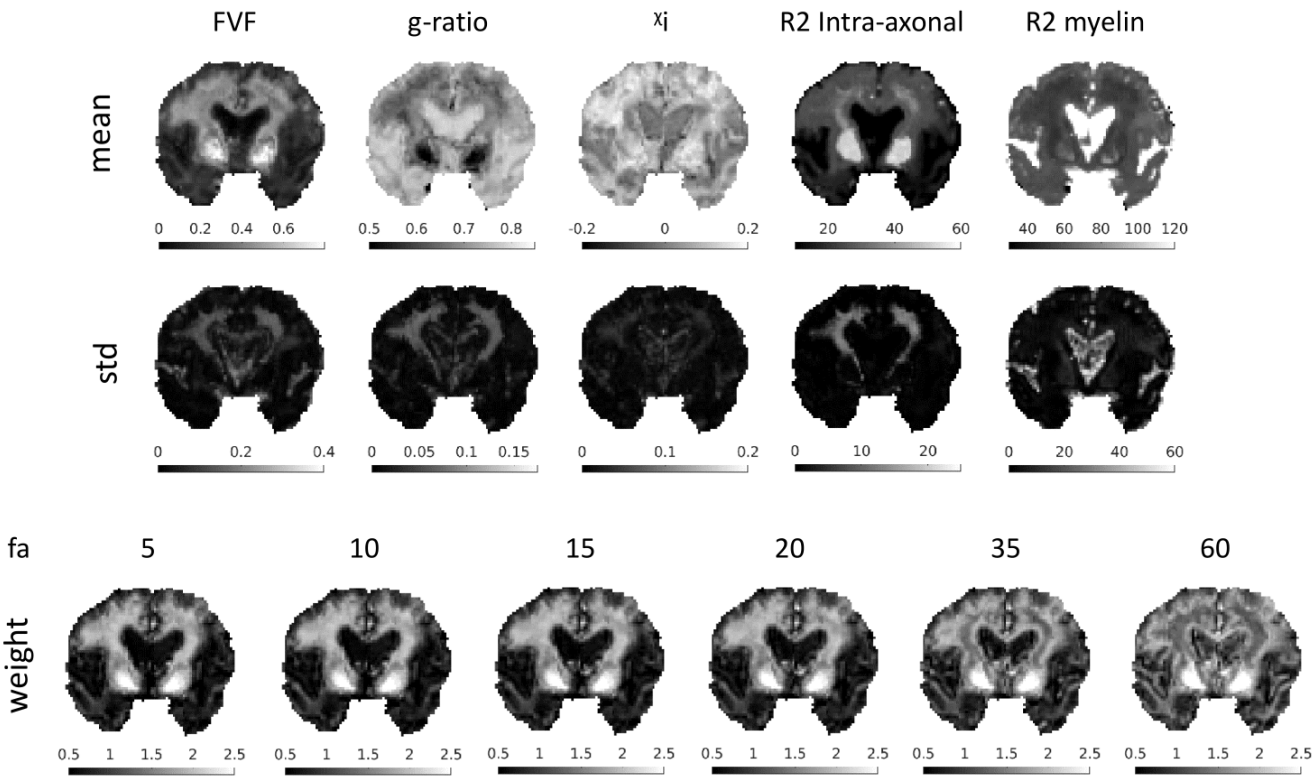


Fig. 9. Top line: Mean parameter maps averaged across the 6 flip angles. Middle line: Corresponding standard deviation maps. Bottom line: Relative water weight for the 6 flip angles.

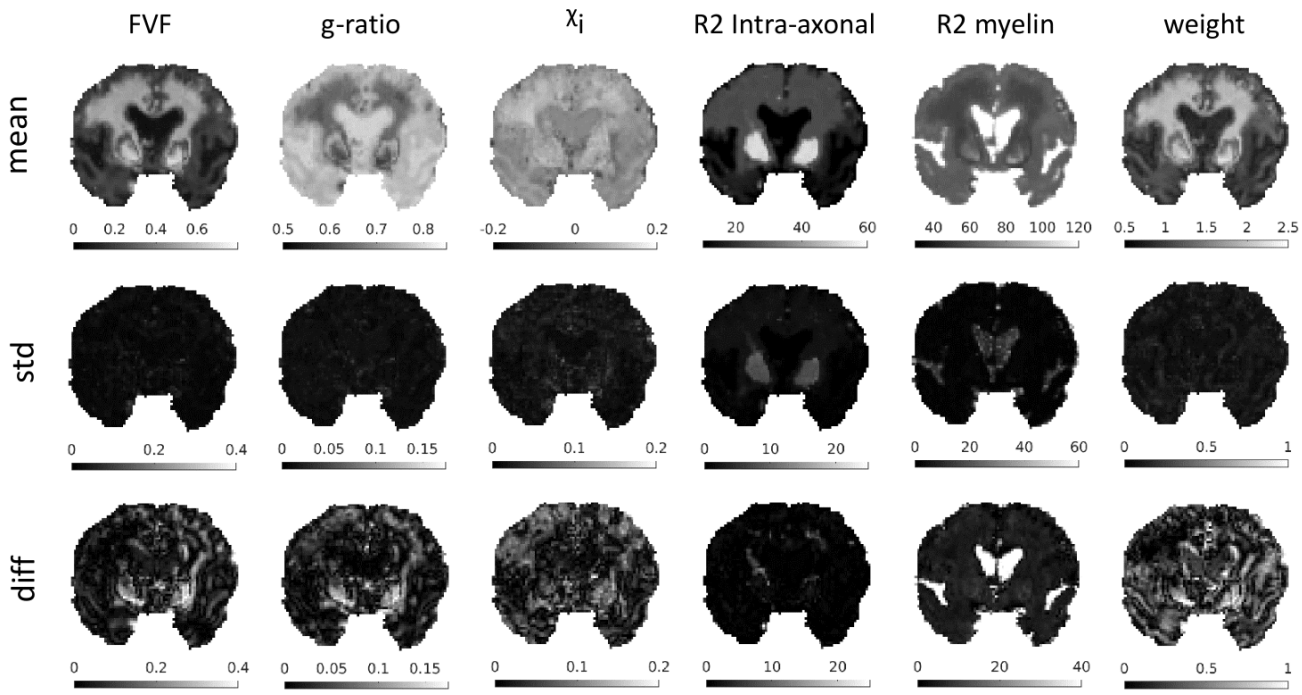


Fig. 10. Top line: Mean parameter maps averaged across 10 combinations of a subset of 3 orientations selected among the 9 original orientations. Bottom line: Corresponding standard deviation maps.

extra axonal T_2^* and the isotropic magnetic susceptibility attributed to the myelin compartment is effectively the difference between the susceptibility of myelin and the free water compartments where there might be ferritin inclusions.

B. Dictionary and deep learning

Many of the simplifications used in our white matter models arise from the need to restrict the number of parameters associated with our network. The size of a dictionary, which in this study had 7 dimensions (see Table I), is around 10 GB, moreover, an increase in the number of variables mapped by the network will result in an increased noise of the parameters estimated. We believe we have restricted the modeling to the most relevant parameters. In particular, we have considered FVF and g-ratio inherent to the model, as described previously the extra-axonal space can have various types of constituents, thus the extra-axonal T_2 cannot be fixed. We choose to free χ_i (allowing this to incorporate magnetic susceptibility in the intra-extra axonal compartment) and to fix χ_a as the major contribution to the magnetic field perturbation comes from the isotropic susceptibility [45]. The compartment water weights were represented by a single variable, the relative water weight that includes the water proton density as well as the degree of T_1-weighting (and chemical exchange) of each compartment. If the myelin sheath is considered having the same properties all other the brain that allows to fix the myelin T2 and release the anisotropic susceptibility χ_a which was reported ranging from -0.15 to -0.09 ppm [15]. A potential direction for future work is to investigate different sets of parameters. For example, one could link the myelin water concentration to the susceptibility of the myelin sheaths by taking into account that the magnetic susceptibility of the phospholipids and water are both known.

Our deep learning network seems robust and systematically converges for each dictionary associated to an experiment with multiple orientations as illustrated in Fig 6. However, extensive manual fine-tuning of the network hyper-parameters was required to achieve this level of agreement. A more systematic approach, while potentially desirable, would need an excessively long computation time. In the future, it may be possible to do this, when access to improved computational resources becomes more common. The *in silico* analysis (see Fig. 7) shows that a dictionary trained with a higher noise level is more robust to noise amplification than a dictionary with matched noise levels. This was attributed to the noise allowing to smear our differences associated with the fact that our "realistic model" produce different signals (see Fig 4) and none of them actually corresponds to the actual white matter mapped. An interesting experiment would be to evaluate the performance of a dictionary including all different noise levels, closer mimicking the signal found in the brain where regions further away from the receiver coils are bound to have a lower SNR. It was observed that the level of noise is within the range that differentiates our 2D models from a real 3D white matter for a relative large range of dispersion values, which effectively makes our network more generalisable.

C. Ex vivo experiment

The human brain scanned on our *ex vivo* experiment was fixed in formalin for 7 months prior to the experiment. It is well known that the microstructural tissue properties change throughout the fixation process, and the final properties of the tissue depend on: the *post mortem* fixation delay, the fixation time, the concentration of formalin and the temperature history [46]–[48]. The T_1 map presented in Fig 8 shows particularly small values revealing a strongly fixed tissues where water has a reduced mobility. This was also clearly visible on the DWI imaging, the mean ADC in white matter being 0.3 $mm^2.s^{-1}$ when a normal *in vivo* values is above 0.8 $mm^2.s^{-1}$ [49]. Fresh tissues do not present such parameter changes and could be an alternative option. However, our current protocol takes 8h without the DWI, in such time window using fresh tissues, would not be sufficiently stable to assume constant microstructural properties over time [50]. Thus, it seems necessary to use fixed human brain, however, the fixation time could be reduced to 6-10 weeks.

The approach presented in this work may find application in the imaging of myelin water with gradient-echo-based acquisitions [8], [51]. Traditionally, myelin water imaging using gradient-echo-based experiments tries to fit 9 independent parameters: three independent signals (separate amplitude, decay rate and frequency shift) for each of the three compartments (intra- and extra- axonal water and myelin) to a ME-GRE signal. The main shortcomings of this approach are that: the model is known to be insufficient (even the simple HCM predicts more complex signal evolution than 3 overlapping exponential signal decays) [25] and the fitting procedure is poorly conditioned. In this work we have shown with simulations that we may obtain acceptable results with as few as 3 orientations (rather than the 9 explored in the *ex vivo* experiment). This may be further improved by including additional diffusion information specific to the intra- and extra-axonal water fractions, given the improved fitting performances obtained recently [52]. Another avenue recently explored is, multi-compartment relaxometry [26], which uses variable flip angle measurements of the ME-GRE signal leveraging the different T_1s of the free water and myelin water compartments to further improve fitting performance. Such frameworks could benefit from realistic WM models, by concatenating along flip angles used instead of being concatenated along rotations. Such an approach would allow to separate water concentration in each compartment from their T_1 weighting giving a more physical meaning to the parameter here dubbed as relative water weight.

D. Ground truth validation

The recovery of microstructural information from the *ex vivo* scans using realistic WM models follows the general expectations for FVF, free water T2, g-ratio and relative water weight. In future experiments, these measurements should be validated by an independent method. One possible avenue is to perform histology on selected excised samples after the scan which could provide a ground-truth, several methods exist to perform such histology analysis. CLARITY is a method using optical 3D imaging combined with a tissue clearing method which can provides neuron density, fiber orientation distribution and cell type classification [53]. X-ray microscopy is an instrument that can generate an entire 3D view of the interior of otherwise optically opaque samples in a non-destructive way [54]. 3D transmission electron microscope (TEM) uses electron as a source of illumination that provides an excellent resolution, better than a classic light microscope [55]. In preliminary work, not shown here, we replicated the fixation process as well as the scanning protocol in a pig brain, from which small sections were extracted to perform 3D TEM analysis. Significant degradation of the myelin sheath for a number of axons was observed, where the myelin sheath appeared unpacked. Such a tissue change would result in a decrease of g-ratio, increase of myelin T2 and proton density, as well as a decrease of χ_i in respect to the *in vivo* case. However, the entire procedure between the brain extraction and the histology was long and could be itself responsible by such alterations.

V. Conclusion

In this paper, we developed an open toolbox [2] to generate 2D white matter models with controlled microstructural properties such as fiber density and variability in the g-ratio using publicly available electron microscopy data. Such models are used to estimate the corresponding field perturbation and derive the multi-echo gradient-echo signals. Although our WM models are limited to 2D, we have demonstrated that they can be satisfactorily used to simulate 3D structures with a relatively high range of dispersion. Finally, dictionaries of GRE signals for 7 different parameters (compartment relaxation values, magnetic susceptibility of myelin, fiber volume fraction and g-ration) associated with white matter properties at a subvoxel level were created were created. This single acquisition dictionaries can then be combined depending on the multiple rotation strategy to create a better conditioned decoding problem and train a deep learning network. We performed several tests to estimate the quality of the subvoxel parameter recovery using our network, depending on the number of sample rotations, echo times used and noise added to the library. Unsurprisingly we found that the network performs better the more data is given as input, thus more rotations and more echoes, but that because of the variations between different white matter modelsit is important to train the network with a level of noise bigger than that of the available data.

The network was demonstrated on an *ex vivo* experiment was performed using a multi-rotation acquisition and provided FVF, g-ratio, T_2 maps clearly revealing brain structures such as the CSF, GM, WM, corpus callosum or globus pallidus. The parameter values (exception for χ_i) follow the expected patterns and were robust for different acquisition protocols.

[2] https://github.com/rhedouin/Whist

VI. Aknowledgements

The authors would like to thank Professor Karla Miller and Dr. Michiel Kleinnijenhuis for providing us the 3D electron microscopy data and its segmentation that was used in the Appendix. This research was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), Grant/Award FOMâĂŘNâĂŘ31/16PR1056that sponsored the positions of Renaud HÃl'douin and Kwok-shing Chan.The authors would also like to acknowledge the fruitful discussions on the topic of this research with Prof. David Norris.

Appendix

A. 3d WM model

To validate the ability of the developed 2D realistic models to describe the 3D structures that exist in a white matter voxel, we compared the signal associated to the 2D models to those of a real 3D WM sample. A segmented [56] 3D EM of the corpus callosum of a mouse was used for this comparison. The resolution of the initial 3D EM dataset was of 7.3x7.3x50 nm, which was subsequently down sampled by a factor of 7 resulting in a quasi isotropic resolution 51x51x50 nm. The FOV of the segmented piece was $20 \times 20 \times 20$ μm (represented on a matrix of $400 \times 400 \times 400$). Using the segmentation 3D EM data, the FVF and g-ratio were computed to be $0,51$ and $0,67$ respectively. Additionally, because the 3D model does not consist of infinitely long structures that are parallel, the fiber dispersion was computed with respect to the average fiber orientation [57], and found to be low $\sigma = 0.04$. In addition to this original model, to study the impact of higher dispersion, 60 axons within the 3D model were selected to create a fiber orientation dispersion of $\sigma = 0.4$. A mask surrounding the selected axons was used to ensure remaining microstructural parameters remained equivalent to those of the the whole sample (FVF $= 0,51$ and g-ratio $= 0,67$). The 3D signal was computed only within the mask and selected axons.

The magnetic susceptibility tensor, X_R, was calculated with respect to the orientation of the phospholipids inside the myelin sheath, using a 3D variant of the process described in the methods section. The obtained tensor map was then used to calculate the magnetic field perturbations in 3D, $\Delta B_0(X(r))$, as described in [30]. Both this processes are straightforward extensions of the 2D case and their implementation is available in our toolbox.

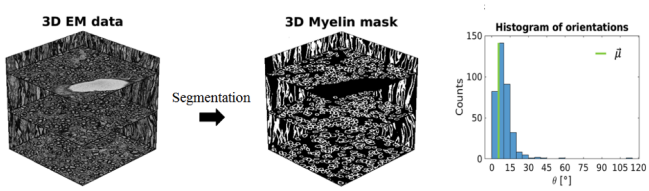


Fig. 11. Raw 3D EM data and myelin segmentation of size 400x400x400. Frequency histogram of the computed axon orientations present in the EM model and the average orientation, $\vec{\mu}$

B. Comparison between 2D and 3D field perturbations

To simulate the fiber dispersion within the 3D samples, an artificial dispersion was introduced into the 2D models by computing the field perturbation for 100 different main magnetic field orientations according to the von-Mises-Fisher distribution [58], the final signal is the sum of signal from 2D models with the 100 different orientations in respect to the main magnetic field.

The 3D models were compared to 10 realistic 2D models, created as described in the methods section, using similar microstructural parameters to those of the 3D samples. Four different different dispersion values ($\sigma = 0, 0.2, 0.4, 0.6$) were simulated. The ME-GRE signals were computed for both 2D and 3D models, with the parameter used in Fig I for TE = 1:1:80 ms. Finally, the 2D and 3D signals were normalized and compared using the root-mean-square-error (RMSE) computed according to:

$$RMSE(\hat{S}_{3D}, \hat{S}_{2D}) = \sqrt{\frac{<(\hat{S}_{3D} - \hat{S}_{2D}), (\hat{S}_{3D} - \hat{S}_{2D})>}{\#TE}} \tag{7}$$

where $<.,.>$ is the complex dot product and $\#TE$ is the number of echos.

Figure 12 shows the signal RMSE between the 2D and 3D models as a function of the orientation of the main magnetic field. In each plot various 2D simulated signals with different dispersion levels are compared to (a) the original 3D model (b) the 3D model with high dispersion. The 2D models with lower dispersion (0 0.2) consistently match that 3D signal with RMSEs below the 2.5%, which is small when taking into account the 4% noise added to the training of our deep learning network. For the high dispersion 3D model (Fig. 12b), the 2D models with high dispersion (0.4 and 0.6) have the lowest RMSE for all magnetic field orientations. When no dispersion is used in the 2D models, the RMSE stays below 5%. The two 3D models considered are best represented with 2D models with similar or slightly higher dispersion values. This finding could be attributed to the additional dispersion associated with each axon that changes direction throughout the 3D model and that is not taken into account in the current dispersion computation.

To conclude, the developed 2D models based on separate library of axons accurately represent a real 3D white matter model. In the future, it could be considered to add

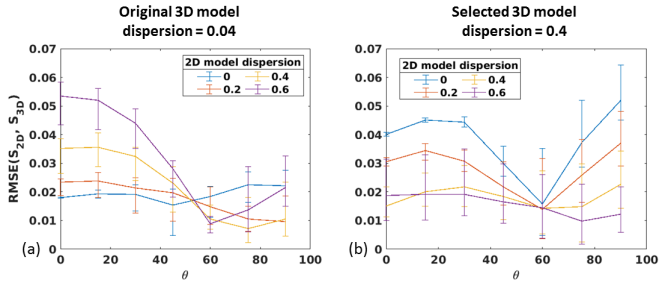


Fig. 12. Plots of the RMSE between the Signal of the 2D models using 4 different dispersion levels and Signal of the 3D models as a function of the orientation of the main magnetic field. In a) the original 3D model with low dispersion (0.04) and in b) the 3D model with high dispersion (0.4)is used as ground truth. The error bars represent the standard deviation across 1o different realistic 2D WM models created

dispersion to the 2D models to better represent a white matter regions with higher dispersion value that could be measured independently with DWI. In ex vivo acquisitions the quality of DTI data is severely hampered (reduced diffusion constant and reduced T2) and from our data it was not possible to apply more advanced diffusion models that can decode this quantity. However, even without dispersion, the RMSE consistently satayed under 5% while 4% noise is added to our dictionary when traing the deep learning network, which suggests that this might not have a large impact.

A situation not considered here and that should have a larger impact are crossing fibers. Fiber dispersion, discussed above, accounts for the spread of the fiber orientations within a bundle of axons while the fiber crossing represents two or more bundles of axons. Significant work on the diffusion community has been devoted to this topic [59]. This could be studied as a future work assuming that such a 3D EM dataset exists.

References

[1] B. Zalc, "The acquisition of myelin: a success story," in *Novartis Foundation symposium*, vol. 276. Wiley Online Library, 2006, p. 15.

[2] S. Love, "Demyelinating diseases," *Journal of clinical pathology*, vol. 59, no. 11, pp. 1151–1159, 2006.

[3] G. Bonnier, B. Maréchal, M. J. Fartaria, P. Falkowskiy, J. P. Marques, S. Simioni, M. Schluep, R. Du Pasquier, J.-P. Thiran, G. Krueger *et al.*, "The combined quantification and interpretation of multiple quantitative magnetic resonance imaging metrics enlightens longitudinal changes compatible with brain repair in relapsing-remitting multiple sclerosis patients," *Frontiers in neurology*, vol. 8, p. 506, 2017.

[4] J. Du, V. Sheth, Q. He, M. Carl, J. Chen, J. Corey-Bloom, and G. M. Bydder, "Measurement of t1 of the ultrashort t2* components in white matter of the brain at 3t," *PloS one*, vol. 9, no. 8, 2014.

[5] Q. He, Y. Ma, S. Fan, H. Shao, V. Sheth, G. M. Bydder, and J. Du, "Direct magnitude and phase imaging of myelin using ultrashort echo time (ute) pulse sequences: A feasibility study," *Magnetic resonance imaging*, vol. 39, pp. 194–199, 2017.

[6] S. D. Wolff and R. S. Balaban, "Magnetization transfer contrast (mtc) and tissue water proton relaxation in vivo," *Magnetic resonance in medicine*, vol. 10, no. 1, pp. 135–144, 1989.

[7] J. G. Sled, "Modelling and interpretation of magnetization transfer imaging in the brain," *Neuroimage*, vol. 182, pp. 128–135, 2018.

[8] E. Alonso-Ortiz, I. R. Levesque, and G. B. Pike, "Multi-gradient-echo myelin water fraction imaging: Comparison to the multi-echo-spin-echo technique," *Magnetic resonance in medicine*, vol. 79, no. 3, pp. 1439–1446, 2018.

[9] K. Susuki, "Myelin: a specialized membrane for cell communication," *Nature Education*, vol. 3, no. 9, p. 59, 2010.

[10] J. H. Duyn and J. Schenck, "Contributions to magnetic susceptibility of brain tissue," *NMR in Biomedicine*, vol. 30, no. 4, p. e3546, 2017.

[11] T. Xu, S. Foxley, M. Kleinnijenhuis, W. C. Chen, and K. L. Miller, "The effect of realistic geometries on the susceptibility-weighted mr signal in white matter," *Magnetic resonance in medicine*, vol. 79, no. 1, pp. 489–500, 2018.

[12] D. A. Yablonskiy and A. L. Sukstanskii, "Biophysical mechanisms of myelin-induced water frequency shifts," *Magnetic resonance in medicine*, vol. 71, no. 6, p. 1956, 2014.

[13] P. Sati, P. van Gelderen, A. C. Silva, D. S. Reich, H. Merkle, J. A. De Zwart, and J. H. Duyn, "Micro-compartment specific t2âĄŐ relaxation in the brain," *Neuroimage*, vol. 77, pp. 268–278, 2013.

[14] S. Wharton and R. Bowtell, "Fiber orientation-dependent white matter contrast in gradient echo mri," *Proceedings of the National Academy of Sciences*, vol. 109, no. 45, pp. 18 559–18 564, 2012.

[15] T. Xu, "Biophysical modeling of white matter in magnetic resonance imaging," Ph.D. dissertation, University of Oxford, 2017.

[16] W. Jung, S. Bollmann, and J. Lee, "Overview of quantitative susceptibility mapping using deep learning–current status, challenges and opportunities," *arXiv preprint arXiv:1912.05410*, 2019.

[17] J. Lee, K. Shmueli, B.-T. Kang, B. Yao, M. Fukunaga, P. Van Gelderen, S. Palumbo, F. Bosetti, A. C. Silva, and J. H. Duyn, "The contribution of myelin to magnetic susceptibility-weighted contrasts in high-field mri of the brain," *Neuroimage*, vol. 59, no. 4, pp. 3967–3975, 2012.

[18] R. Gil, D. Khabipova, M. Zwiers, T. Hilbert, T. Kober, and J. P. Marques, "An in vivo study of the orientation-dependent and independent components of transverse relaxation rates in white matter," *NMR in Biomedicine*, vol. 29, no. 12, pp. 1780–1790, 2016.

[19] J. Lee, H.-G. Shin, W. Jung, Y. Nam, S.-H. Oh, and J. Lee, "An r2* model of white matter for fiber orientation and myelin concentration," *Neuroimage*, vol. 162, pp. 269–275, 2017.

[20] S. Papazoglou, T. Streubel, M. Ashtarayeh, K. J. Pine, L. J. Edwards, M. Brammerloh, E. Kirilina, M. Morawski, C. Jäger, S. Geyer *et al.*, "Biophysically motivated efficient estimation of the spatially isotropic component from a single gradient-recalled echo measurement," *Magnetic resonance in medicine*, vol. 82, no. 5, pp. 1804–1811, 2019.

[21] Y. Nam, J. Lee, D. Hwang, and D.-H. Kim, "Improved estimation of myelin water fraction using complex model fitting," *NeuroImage*, vol. 116, pp. 214–221, 2015.

[22] P. Van Gelderen, J. A. De Zwart, J. Lee, P. Sati, D. S. Reich, and J. H. Duyn, "Nonexponential t2* decay in white matter," *Magnetic resonance in medicine*, vol. 67, no. 1, pp. 110–117, 2012.

[23] E. Alonso-Ortiz, I. R. Levesque, and G. B. Pike, "Impact of magnetic susceptibility anisotropy at 3 t and 7 t on t2*-based myelin water fraction imaging," *NeuroImage*, vol. 182, pp. 370–378, 2018.

[24] J. Lee, Y. Nam, J. Y. Choi, E. Y. Kim, S.-H. Oh, and D.-H. Kim, "Mechanisms of t2* anisotropy and gradient echo myelin water imaging," *NMR in Biomedicine*, vol. 30, no. 4, p. e3513, 2017.

[25] S. Wharton and R. Bowtell, "Gradient echo based fiber orientation mapping using r2* and frequency difference measurements," *Neuroimage*, vol. 83, pp. 1011–1023, 2013.

[26] K.-S. Chan and J. P. Marques, "Brain tissue multi-compartment relaxometry - an improved method for in vivo myelin water imaging," in *ISMRM*, 2019.

[27] A. Zaimi, T. Duval, A. Gasecka, D. Côté, N. Stikov, and J. Cohen-Adad, "Axonseg: open source software for axon and myelin segmentation and morphometric analysis," *Frontiers in neuroinformatics*, vol. 10, p. 37, 2016.

[28] T. Mingasson, T. Duval, N. Stikov, and J. Cohen-Adad, "Axon-packing: an open-source software to simulate arrangements of axons in white matter," *Frontiers in neuroinformatics*, vol. 11, p. 5, 2017.

[29] F. Sepehrband, K. A. Clark, J. F. Ullmann, N. D. Kurniawan, G. Leanage, D. C. Reutens, and Z. Yang, "Brain tissue compartment density estimated using diffusion-weighted mri yields tissue parameters consistent with histology," *Human brain mapping*, vol. 36, no. 9, pp. 3687–3702, 2015.

[30] W. Li, C. Liu, T. Q. Duong, P. C. van Zijl, and X. Li, "Susceptibility tensor imaging (sti) of the brain," *NMR in Biomedicine*, vol. 30, no. 4, p. e3540, 2017.

[31] S. W. Choy, E. Bagarinao, H. Watanabe, E. T. W. Ho, S. Maesawa, D. Mori, K. Hara, K. Kawabata, N. Yoneyama, R. Ohdake *et al.*, "Changes in white matter fiber density and morphology across the adult lifespan: A cross-sectional fixel-based analysis," *Human Brain Mapping*, 2020.

[32] S. Mohammadi, D. Carey, F. Dick, J. Diedrichsen, M. I. Sereno, M. Reisert, M. F. Callaghan, and N. Weiskopf, "Whole-brain in-vivo measurements of the axonal g-ratio in a group of 37 healthy volunteers," *Frontiers in neuroscience*, vol. 9, p. 441, 2015.

[33] J. Mollink, M. Kleinnijenhuis, A.-M. v. C. van Walsum, S. N. Sotiropoulos, M. Cottaar, C. Mirfin, M. P. Heinrich, M. Jenkinson, M. Pallebage-Gamarallage, O. Ansorge *et al.*, "Evaluating fibre orientation dispersion in white matter: Comparison of diffusion mri, histology and polarized light imaging," *Neuroimage*, vol. 157, pp. 561–574, 2017.

[34] K. Ginsburger, F. Poupon, J. Beaujoin, D. Estournet, F. Matuschke, J.-F. Mangin, M. Axer, and C. Poupon, "Improving the realism of white matter numerical phantoms: a step toward a better understanding of the influence of structural disorders in diffusion mri," *Frontiers in Physics*, vol. 6, p. 12, 2018.

[35] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. L. Sunshine, J. L. Duerk, and M. A. Griswold, "Magnetic resonance fingerprinting," *Nature*, vol. 495, no. 7440, pp. 187–192, 2013.

[36] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[37] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[38] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 56, no. 2, pp. 411–421, 2006.

[39] X. Li, D. S. Vikram, I. A. L. Lim, C. K. Jones, J. A. Farrell, and P. C. van Zijl, "Mapping magnetic susceptibility anisotropies of white matter in vivo in the human brain at 7 t," *Neuroimage*, vol. 62, no. 1, pp. 314–330, 2012.

[40] D. A. Yablonskiy, "Quantitation of intrinsic magnetic susceptibility-related effects in a tissue matrix. phantom study," *Magnetic resonance in medicine*, vol. 39, no. 3, pp. 417–428, 1998.

[41] D. A. Yablonskiy and A. L. Sukstanskii, "Lorentzian effects in magnetic susceptibility mapping of anisotropic biological tissues," *Journal of Magnetic Resonance*, vol. 292, pp. 129–136, 2018.

[42] P. van Gelderen and J. H. Duyn, "White matter intercompartmental water exchange rates determined from detailed modeling of the myelin sheath," *Magnetic resonance in medicine*, vol. 81, no. 1, pp. 628–638, 2019.

[43] A. L. Sukstanskii, J. Wen, A. H. Cross, and D. A. Yablonskiy, "Simultaneous multi-angular relaxometry of tissue with mri (smart mri): Theoretical background and proof of concept," *Magnetic resonance in medicine*, vol. 77, no. 3, pp. 1296–1306, 2017.

[44] D. J. Piñero and J. R. Connor, "Iron in the brain: an important contributor in normal and diseased states," *The Neuroscientist*, vol. 6, no. 6, pp. 435–453, 2000.

[45] W. Li, B. Wu, A. V. Avram, and C. Liu, "Magnetic susceptibility anisotropy of human brain in vivo and its molecular underpinnings," *Neuroimage*, vol. 59, no. 3, pp. 2088–2097, 2012.

[46] A. S. Shatil, M. N. Uddin, K. M. Matsuda, and C. R. Figley, "Quantitative ex vivo mri changes due to progressive formalin fixation in whole human brain specimens: longitudinal characterization of diffusion, relaxometry, and myelin water fraction measurements at 3t," *Frontiers in medicine*, vol. 5, p. 31, 2018.

[47] C. Birkl, M. Soellradl, A. M. Toeglhofer, S. Krassnig, M. Leoni, L. Pirpamer, T. Vorauer, H. Krenn, J. Haybaeck, F. Fazekas *et al.*, "Effects of concentration and vendor specific composition of formalin on postmortem mri of the human brain," *Magnetic resonance in medicine*, vol. 79, no. 2, pp. 1111–1115, 2018.

[48] T. M. Shepherd, J. J. Flint, P. E. Thelwall, G. J. Stanisz, T. H. Mareci, A. T. Yachnis, and S. J. Blackband, "Postmortem interval alters the water relaxation and diffusion properties of rat nervous tissueâĂŤimplications for mri studies of human autopsy samples," *Neuroimage*, vol. 44, no. 3, pp. 820–826, 2009.

[49] R. Sener, "Diffusion mri: apparent diffusion coefficient (adc) values in the normal brain and a classification of brain disorders based on adc values," *Computerized medical imaging and graphics*, vol. 25, no. 4, pp. 299–326, 2001.

[50] F. A. Duck, *Physical properties of tissues: a comprehensive reference book*. Academic press, 2013.

[51] L. E. Lee, E. Ljungberg, D. Shin, C. R. Figley, I. M. Vavasour, A. Rauscher, J. Cohen-Adad, D. K. Li, A. L. Traboulsee, A. L. MacKay *et al.*, "Inter-vendor reproducibility of myelin water imaging using a 3d gradient and spin echo sequence," *Frontiers in neuroscience*, vol. 12, p. 854, 2018.

[52] K.-S. Chan and J. P. Marques, "Diffusion-weighted imaging informed gradient-echo myelin water imaging," in *ESMRMB*, 2019.

[53] M. Morawski, E. Kirilina, N. Scherf, C. Jäger, K. Reimann, R. Trampel, F. Gavriilidis, S. Geyer, B. Biedermann, T. Arendt *et al.*, "Developing 3d microscopy with clarity on human brain tissue: Towards a tool for informing and validating mri-based histology," *Neuroimage*, vol. 182, pp. 417–428, 2018.

[54] S. R. Stock, *Microcomputed tomography: methodology and applications*. CRC press, 2019.

[55] N. De Jonge, R. Sougrat, B. M. Northan, and S. J. Pennycook, "Three-dimensional scanning transmission electron microscopy of biological specimens," *Microscopy and Microanalysis*, vol. 16, no. 1, pp. 54–63, 2010.

[56] M. Kleinnijenhuis et al., "A 3d electron microscopy segmentation pipeline for hyper-realistic diffusion simulations," 2017.

[57] (2019) Fsl - dispersion calculation. Analysis Group, FMRIB, Oxford, UK. [Online]. Available: https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide?action=AttachFile&do=view&target=dyads.jpg

[58] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra, "Clustering on the unit hypersphere using von mises-fisher distributions," *Journal of Machine Learning Research*, vol. 6, no. Sep, pp. 1345–1382, 2005.

[59] H. Farooq, J. Xu, J. W. Nam, D. F. Keefe, E. Yacoub, T. Georgiou, and C. Lenglet, "Microstructure imaging of crossing (mix) white matter fibers from diffusion mri," *Scientific reports*, vol. 6, p. 38927, 2016.