Axon Diameter Measurements using Diffusion MRI are Infeasible

Michael Paquette[1], Cornelius Eichner[1], Thomas R. Knösche[1], and Alfred Anwander[1]

[1]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Abstract

The feasibility of non-invasive axonal diameter quantification with diffusion MRI is a strongly debated topic due to the neuroscientific potential of such information and its relevance for the axonal signal transmission speed. It has been shown that under ideal conditions, the minimal diameter producing detectable signal decay is bigger than most human axons in the brain, even using the strongest currently available MRI systems. We show that resolving the simplest situations including multiple diameters is unfeasible even with diameters much bigger than the diameter limit. Additionally, the recently proposed effective diameter resulting from fitting a single value over a distribution is almost exclusively influenced by the biggest axons. We show how impractical this metric is for comparing different distributions. Overall, axon diameters currently cannot be quantified by diffusion MRI in any relevant way.

1 Introduction

In-vivo estimation of axon diameters has been an important goal of many researchers since the inception of diffusion MRI. As the diameter of a myelinated axon is one of the main determiner of its signal transmission velocity [30, 13], the availability of this structural information would greatly facilitate description and functional modeling of the brain communication pathways on an individual basis [32]. Detailed knowledge of tract specific axonal diameters would provide insight into detailed and mechanistic relationships between brain structure and important aspects of brain function, including development and learning. The capacity of dMRI to noninvasively probe cellular and axonal boundaries at the micrometer level seemed a promising method to pursue this aim.

The impact of restricted incoherent motion of water molecules on diffusion weighted NMR signals has already been described in the early days of MR spectroscopy [26, 37]. However, these models only describe the diffusion process happening in the perpendicular cross-section of the axon. Using them to approximate axonal diameters requires prior knowledge of the tissue orientations, an equal diameter of all axons in the probed volume, as well as the absence of extra-axonal signals. A common strategy to bring it to the in-vivo 3D acquisition setting has been to combine one or many cylindrical compartments, describing the intra-axonal diffusion, with additional compartments describing the extra-axonal Gaussian diffusion process [4, 5, 2, 10].

Despite the tremendous overestimation of axonal diameters arising from the use of multi compartment models [16] compared to electron microscopy ground truth [1, 24], these models are still seen as promising by a part of the community. This dilemma can be attributed to the fact that the relative *trend* of fitted diameters was argued to be somewhat plausible across the different parts of the corpus callosum [16] and that multi-compartment models in dMRI are difficult to fit reliably as they are essentially weighted sums of exponential functions.

Recent work highlighted an unavoidable sensitivity issue for detecting axon diameters of realistic size in the human brain, even with the latest high-end MRI systems [14, 27]. It proposes an "axon diameter limit" ($d_{\min}$) which corresponds to the smallest diameter that can be differentiated from a stick of diameter zero for given sequence parameters under ideal conditions. This $d_{\min}$ is computed from the most generous setting and is therefore a lower bound on the unbiased smallest diameter detectable for data deviating from the idealized case of diffusion signal arising only from parallel cylinders of equal diameter. The diameter limit suggests that previous "trends" in the estimated diameters are not supported by the measured data. Indeed, not only is the expected signal decay for restricted diffusion in realistic human axons size very small, it is also insensitive to changes in the gradient spacing (Δ), which is typically the parameter varied when the "small-big-small diameter trend" of the corpus callosum is observed. The large signal decay observed could be caused by noise, errors in the compartment separation or by other types of time-dependent diffusion such as diffusion signal from the extra-axonal compartment, which is sensitive to Δ.

In this work, we employ extensive simulations of restricted diffusion MRI measurements under optimal conditions to concretely showcase the limitations of axon diameter mapping. We first show the sensitivity of MR signals to axon diameters. Secondly, we show the axon diameter limit in action in the case of fitting a single diameter. We then show the unresolvability of extending from single diameters to estimations of distributions of axonal diameters. Finally, we highlight the difficulty of interpreting a single diameter fitted over a distribution (so-called effective diameter [8, 38]), even with a complete understanding of the averaging mechanisms

projecting the distribution on this single value. The limitations of MR axon diameter estimation discussed in this work affect every diffusion model fitting diameters, as they all rely on the capacity of dMRI to detect small signal decay from inside cylinder compartment [5, 2, 18, 10, 17].

2 Methods

2.1 Relevant parameters

Throughout this work, we used numerical simulations to showcase the sensitivity of dMRI to axon diameters. It is therefore crucial to use realistic values for the various physical parameters. We describe each parameter, their realistic ranges, and our default choices. Particularly, we are concerned with the order of magnitude of the quantities and their scaling behavior (see eq. 1). For completion, we provide scripts to recompute any quantity, figure or experiment, for any choice of parameters (`https://github.com/mpaquette/axDiamFig`).

Axon (cylinder) diameter (d): The smaller the diameter, the smaller is the maximal displacement of the water molecules, as we assume impermeable axonal walls. This restricted water diffusion perpendicular to the axon will induce a small signal change proportional to the mean squared displacement inside the circular cross-section. Prior results from histological assessments show that human axons in the white matter of the brain have diameters in the order of 1 μm [1, 24]. Typical distributions of diameters tend to peak around 0.5-1.0 μm with maximum axon diameters around 2.5-5 μm (see fig 4). Informally, the minimum sensitivity required to properly qualify such distributions has to be smaller than the peak of the distribution.

Unrestricted diffusivity of the medium (D_0): The lower the diffusivity is, the more time it takes for the diffusion process to saturate inside of the restricted compartment. The literature reports diffusivity values between 1.5 $\mu\text{m}^2/\text{ms}$ and 2.7 $\mu\text{m}^2/\text{ms}$ for the in-vivo intra-axonal water compartment [40, 12]. In the case of post-mortem measurements, both the reduced tissue temperature and the fixation process reduce the tissue diffusivity [29]. Reported values for post-mortem diffusivities are around 1/3 - 1/4 of that of in-vivo [15]. In our simulations we assume the following diffusivities: $D_{0,in-vivo} = 2\ \mu\text{m}^2/\text{ms}$ ($2 \times 10^{-9}\ \text{m}^2/\text{s}$), and $D_{0,post-mortem} = 0.66\ \mu\text{m}^2/\text{ms}$ ($0.66 \times 10^{-9}\ \text{m}^2/\text{s}$).

Diffusion gradient magnitude (G): The diffusion gradient hardware varies among the different types of MRI scanners. The strength of diffusion gradient directly affects the signal decay in diffusion, *i.e.* the same total displacement of water molecule produces a bigger signal decay with a stronger gradient. Typical clinical scanners tend to have weaker gradients ($G_{\max} = 40$ mT/m), while gradient coils in preclinical small-bore scanners can produce magnetic field gradients as strong as 1500 mT/m. For human in-vivo measurements, the Siemens Connectom MRI scanner (Siemens

Healthineers, Erlangen, Germany) is the system which produces by far the strongest diffusion gradients ($G_{\max} = 300$ mT/m). In our simulations we use $G = 300$ mT/m, as one of the goals associated with the development of this specific MRI system was to enable in-vivo axon diameter estimation.

Diffusion gradient duration (δ): In the relevant regimes for human axon diameter estimation, the duration of the diffusion gradient pulse δ is the parameter probing the time-dependent diffusivity of restricted diffusion. Indeed, with a shortest achievable gradient duration around 5 ms on a human MRI system, we are well into the regime where the gradient duration is comparable with the saturation time of the restricted compartment. In this regime, longer gradient pulses increase sensitivity (see sec. A.1). We limit the simulations to $\delta_{\max} = 40$ ms as longer pulses are impractical, as they increase the echo times of the acquisition, resulting in additional signal losses.

Diffusion gradient separation (Δ): In the relevant regimes for human axon diameter estimation, the diffusion process is already saturated during the gradient application and varying the separation of the diffusion gradient pulses Δ provides no extra sensitivity to restricted diffusion (see sec. A.1 and fig. 7). Therefore, to maximize signal, we use $\Delta = \delta$. In practice, varying Δ *could* still be necessary for multi-compartment models where it is necessary to disentangle intra- and extra-axonal signal contributions.

Signal to noise ratio (SNR): Ultimately, the SNR is the key parameter upon which "sensitivity" is defined. Throughout the simulations presented in this study, we corrupt signals with Gaussian noise (for simplicity and to produce a best case scenario), *i.e.* $S_{noisy} = S_{noiseless} + \epsilon$ where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Since we only look at idealized diffusion effects, our signals have value of 1 at the b0 (no diffusion gradient applied), and therefore the SNR is defined as SNR $= \sigma^{-1}$. For comparison, the SNR of the b0 in the corpus callosum for a single in-vivo volume on the Connectom system with echo time of 70 ms, repetition time 7500 ms and resolution 1.8 mm isotropic is around 20. We showcase results for SNR = 30 and some results for SNR = 300, which correspond to 100 averages of a high quality Connectom acquisition, or parameter estimations. Some diameter estimation approaches use aggregated fitting strategies such as region of interest (ROI) averaging or averaging along a tractography streamline path [11, 7, 6, 3] to increase the nominal SNR. Obviously, these aggregated strategies make strong assumptions on tissue composition and orientation homogeneity in a region or along the entire pathway. It is unclear if the SNR gains of such strategies outweigh the bias due to tissue inhomogeneities in neighboring voxels as these methods still suffer from diameter overestimation [6].

2.2 dMRI signal sensitivity to diameter

Diffusion MRI contrast is related to the bulk displacement of the water molecules during the diffusion encoding, which causes the measured signal decay. Inside restricted compartments such as the cross-section of a cylinder, the maximal displacement is capped by the boundary, potentially producing much smaller signal decays than produced by free diffusion. These restricted diffusion processes can be classified into different time regimes. On short time scales, the bulk of water molecules has not yet interacted with the boundary, and therefore behaves as in free diffusion. In the long time regime, most molecules have significantly interacted with the boundary and their position at any given time does not correlate with their initial position inside the cross section of the axon; *i.e.* the signal as reached maximal decay.

The general perpendicular signal decay formula for a cylinder using a Pulsed Gradient Spin Echo (PGSE) diffusion sequence [35] was first described by Neuman [26] and then extended by Van Gelderen [37] to account for cases where $\Delta \neq \delta$ (eq. 3). For the parameter ranges described in sec. 2.1, the Neuman long time limit (eq. 4) produce almost indistinguishable results. In this work, we use eq. 3 truncated to 50 terms to generate and fit signals arising from restricted diffusion.

For realistic acquisition and biological relevant parameter values (see sec 2.1), the diffusion process falls into the long time regime and the expected signal decay is small compared to noise amplitude at typical SNR. Using eq. 3, we simulated the expected MR signal decay for a multitude of combinations and we report the **decay percentage** values in Table 1. To cover a wide range of biological, experimental and instrumental parameters, we simulated restricted diffusion MRI signals using (i) both in-vivo and post-mortem diffusivities, (ii) clinical gradient systems and high-end Connectom gradients, and (iii) small to large human axons diameter.

Our simulations indicated that dMRI is not very sensitive to the axonal diameter in realistic situations. For example, using optimal in-vivo setting (Connectom strength gradients, very long diffusion pulse and in-vivo diffusivity) for an axon diameter of 1 micrometer the process only produces a "contrast" of 0.12% signal decay which is equal to one standard deviation of Gaussian noise with $SNR \approx 833$. To be able to statistically identify this signal decay, we would typically need a decay to be at least bigger than $\sim$ 2 standard deviation of the noise, depending on the choice of the significance level. To reach such a low noise level would require SNR $\approx$ 1667. Hence, for realistic SNRs, small diameters cannot be differentiated from the noise level in the image.

2.3 Axon diameter limit

To formalize the notion of sensitivity into a workable form using signal decay and SNR, Nilsson et al. [27] introduced the diameter resolution limit ($d_{\min}$). It is defined as the smallest diameter such that the MR signal decay can be statistically differ-

Acquisition parameters		**In-vivo** ($D_0 = 2.0\ \mu\text{m}^2/\text{ms}$)		
$\delta = \Delta$ (ms)	G (mT/m)	$d = 0.5\ \mu$m	$d = 1.0\ \mu$m	$d = 2.0\ \mu$m
10	40	3.2×10^{-5}	5.2×10^{-4}	8.1×10^{-3}
40	40	1.3×10^{-4}	2.1×10^{-3}	3.3×10^{-2}
10	300	1.8×10^{-3}	2.9×10^{-2}	4.6×10^{-1}
40	300	7.3×10^{-3}	1.2×10^{-1}	1.8
Acquisition parameters		**Post-mortem** ($D_0 = 0.66\ \mu\text{m}^2/\text{ms}$)		
$\delta = \Delta$ (ms)	G (mT/m)	$d = 0.5\ \mu$m	$d = 1.0\ \mu$m	$d = 2.0\ \mu$m
10	40	9.8×10^{-5}	1.6×10^{-3}	2.4×10^{-2}
40	40	3.9×10^{-4}	6.3×10^{-3}	9.9×10^{-2}
10	300	5.5×10^{-3}	8.7×10^{-2}	1.3
40	300	2.2×10^{-2}	3.5×10^{-1}	5.4

Table 1: MR signal **decay** (in percent) for various diffusivities, acquisition parameters and axon diameters. We note that if we have SNR = 30, a noise realization of one standard deviation has a magnitude 3.3% signal decay. This showcases the difficulty of detecting and differentiating the signal decay caused by different diameter. For the post-mortem case, using the somewhat big $d = 1\ \mu$m and strong Connectom-like acquisition (G = 300 mT/m), we are expecting a signal decay of 0.35%. To be able to statistically identify this signal decay, we would typically need a decay to be at least bigger than ~ 2 standard deviation of the noise (depending on choice of significance level), which would require SNR ≈ 571.

entiated from no decay (in the limiting case $d \rightarrow 0$) for a given SNR and choice of significance level for the Z-test (α). The decay limit is given by $\bar{\sigma} = Z_{1-\alpha}/\text{SNR}$. We use eq. 1 to find d_{min} corresponding to the decay limit. We use $\alpha = 0.05$ ($Z_{1-0.05} = 1.645$) for the entirety of this work.

$$d_{\text{min}} = \left(\frac{768}{7}\frac{\bar{\sigma}D_0}{\gamma^2\delta G^2}\right)^{1/4} \overset{(\alpha\,=\,0.05)}{=} \left(\frac{4512}{25}\frac{D_0}{\gamma^2\delta G^2\text{SNR}}\right)^{1/4} \tag{1}$$

Practically, the main implications of this framework are governed by the exponents of the individual parameters. We can see for instance that halving the diameter limit requires 4-fold increase in gradient strength or 16-fold increase in SNR ($\sim$ 256 repetitions averaged). Table 2 showcases some values of d_{min} for in-vivo and post-mortem diffusivities, a long gradient pulse, various gradient strengths (clinical, Connectom, and small-bore preclinical) for various SNRs. We see that even in the idealized case [27], we obtain $d_{\text{min}} = 2.56$ μm for the in-vivo Connectom case at realistic SNR, falling quite short of our minimum target of around 1 μm. At SNR = 164 ($\sim$ 5 times higher than baseline, $\sim$ 25 averages), we have 1.77 μm. In this example, we need tissue with low post-mortem diffusivity and ultra-strong gradients of the strongest preclinical scanner (G = 1500 mT/m) to reach the initial goal of $d_{\text{min}} \leq 1$ μm, showcasing the practical limitations arising from the fourth root scaling in eq. 1.

Parameters			**SNR**		
D_0 (μm^2/ms)	$\delta = \Delta$ (ms)	G (mT/m)	**164**	**65.6**	**32.8**
2.0	40	40	4.69	5.89	7.01
2.0	40	300	1.71	2.15	2.56
2.0	40	1500	0.77	0.96	1.14
0.66	40	40	3.55	4.47	5.31
0.66	40	300	1.30	1.63	1.94
0.66	40	1500	0.58	0.73	0.87

Table 2: Values of d_{min} (μm) (eq. 1) for various parameters at significance level $\alpha = 0.05$ (*i.e.* signal decay stronger than 1.645 standard deviations of noise distribution). The selected SNRs (164, 65.6, 32.8) correspond to minimum detectable signal decays of 1%, 2.5% and 5%.

To visualise the impact of d_{min}, we plot the spread of recovered diameters in fig. 1. For each diameter between 0.1 μm and 5 μm, we generated 10000 noisy restricted signals and added Gaussian noise with SNR 30 and 300. The signals are generated for realistic in-vivo settings ($D_0 = 2$ μm^2/ms) with a Connectom-like acquisition (single "direction/average", G = 300 mT/m, $\delta = \Delta = 40$ ms). The different SNRs are scaled copies and we see that the mean recovered diameter is biased for diameters smaller than d_{min}. The bias occurs because the average detected diameters become

driven by the signal decay corresponding to one standard deviation of noise. Hence, the result suffers not only from uncertainty, but also from systematic bias.

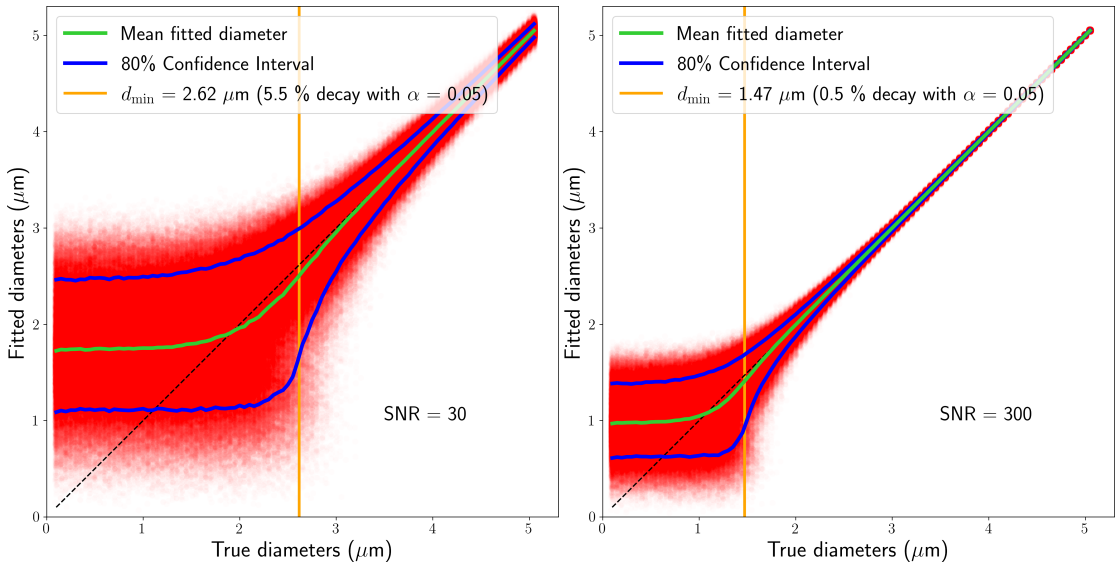

Figure 1: Scatter plot of fitted diameters with mean fitted diameter (green line) and 80% confidence interval (blue lines). For each diameter between 0.1 μm and 5 μm, we generated 10000 noisy restricted signals using eq. 1 and Gaussian noise of SNR 30 (left) and 300 (right). The signals are generated for realistic in-vivo setting ($D_0 = 2$ μm^2/ms) with a Connectom-like acquisition (single "direction", G = 300 mT/m, $\delta = \Delta = 40$ ms). The orange line corresponds to $d_{\min}$ using the framework by [27].

It is necessary to insist on what the definition of $d_{\min}$ truly implies, because it is often misunderstood as being the diameter above which fitting will be stable. The formalism of this section is a way to calculate the smallest signal decay *difference* which is statistically differentiable from 0. We can assess if the SNR and acquisition parameters are enough to differentiate two arbitrary diameters, by verifying that their produced signal decay difference is bigger than $\bar{\sigma}$. If we set one of those diameters to 0 and we look for the smallest second diameter above the threshold, we get $d_{\min}$. The minimum diameter only assures us that the distribution of a noisy signal decay around a true signal decay from a diameter bigger than $d_{\min}$ doesn't "overlap significantly" with a signal decay of 0 (*i.e.* less than α of the distribution is below 0).

2.4 Axon diameter distributions

In previous sections, we focused on the sensitivity of dMRI for axon populations of a single diameter within a voxel. However, the white matter is composed of axons with multiple diameters spanning a large range [1, 24] . Therefore, it is sensible to fit

a full distribution of diameters to the measured signal. This strategy can be implemented in multiple ways, such as enforcing a parameterized distribution family such as a gamma distribution over the relative axon counts, fitting volume fractions for a binned discretized distribution or by fitting multiple cylinder compartments with diameters as a free parameter. Intuitively, moving from single diameter estimations to any type of distribution will increase the $d_{\min}$, because adding additional degrees of freedom to a model increases the variance of the fitted parameters [19]. However, the fitting of axonal diameter distributions to dMRI signals is plagued by more than a simple increase to the related $d_{\min}$.

In this chapter we show that even the simplest model with multiple diameters has infinitely many completely different solutions for realistic parameters (sec. 2.1). These simulations suggest that any "trend" of different diameters seen in images using such models is not supported by theory and is likely driven by either the regularization terms in the fit or by an effect unrelated to diameter, like noise, errors in the compartment separation or by other types of time-dependent diffusion such as a diffusion signal from the extra-axonal compartment.

When we describe distributions of axon diameters, $P_{\text{axon}}(d)$, we refer to distributions over the number (*axon count*) of axons for each diameter inside a voxel. Under the assumption that axons of different diameter have the same proton density, the *spin count* distribution becomes a cylinder volume-weighting of the *axon count* distribution, $P_{\text{spin}}(d) = P_{\text{axon}}(d)\frac{\text{Vol}(d)}{\int \text{Vol}(d')\text{d}d'}$. Since the different axons are implicitly assumed to be of the same length inside the voxel, the volume-weighting becomes a cross-section area-weighting ($P_{\text{spin}}(d) = P_{\text{axon}}(d)\frac{d^2}{\int d'^2 \text{d}d'}$). The normalized spin counts are also often referred to as the **volume fractions** of each axon diameter, representing the relative volume of water inside the axons of a given diameter. When the water molecules inside the axons of different diameters have the same magnetic properties (*i.e.* identical T_2, T_1, etc), the **signal fractions** are equivalent to the volume-normalized *axon count* distribution. In this study, the conversion between volume and signal fraction only depends on cross-sectional area re-weighting.

In this experiment, we define the simplest distribution, a signal generated from a population of two parallel very big axon diameters in roughly equal proportion (with **signal fractions**: 30% $d_1 = 4.5$ μm and 70% $d_2 = 3.5$ μm, equivalent to **volume fractions** of 41.5% and 58.5%) (fig. 2). We then plot the mean absolute difference between this (noiseless) signal and the signals generated for all the other possible configurations.

Similarly to how we only used a single "acquisition" (with maximally sensitive Connectom-like parameters) for the single parameters estimation in fig. 1, here we use Connectom-like acquisition parameters with three different gradient pulse durations to mimic the minimal requirements of uniquely fitting a three parameter model (two diameters and one signal fraction). The acquisition parameters were selected such that they provide sensitivity (long δ) and that the biggest individual $d_{\min}$

is comfortably below the smallest diameter in the ground truth (G = 300 mT/m, $\Delta = 50$ ms, $\delta = [30, 40, 50]$ ms). This two-cylinder model has a three dimensional space of possible parameter configurations: the first diameter, the second diameter and the signal fraction (of the first cylinder). In fig. 2, the parameter space is sliced in the signal fraction direction every 5% and shown as a sequence of 2D plots spanning all pairs of diameters. Regions of solid colors across all slices correspond to regions of the parameter space producing similar signal decay in this noiseless setting. For instance, the blue region corresponds to configurations producing a signal with less than 1% signal decay difference from the ground truth, making them indistinguishable at regular SNR (for example, 1% signal decay correspond to SNR = 164 for significance level $\alpha = 0.05$). The blue region spans a surface across many unrelated pairs of diameters and signal fractions, showcasing the unresolvability of the simplistic two-diameter distribution under optimal conditions (ground truth perfectly matching the model and no other compartments to disentangle). The axon population diameters were chosen to be very big to highlight the fundamental problem of distribution fitting, for similar figures with smaller diameters, see Sec. A.4 where the effect is amplified.

In fig. 3, we repeat the previous experiment with gamma distributed axon diameter counts instead of the two-diameter distribution. We generated a signal using a population of cylinders where the *count* for each diameter follows a gamma distribution (shape = 2.25 and scale = 0.4 with peak at 0.5 μm) using the same diffusivities and acquisition parameters as in fig. 2. We show the mean absolute difference between our (noiseless) signal and signal generated from gamma distributions spanning shapes up to 9 and peak location up to 3 μm. We note that a gamma distribution $\Gamma(k, \theta)$ of shape k and scale θ has its peak at $(k-1)\theta$ for $k \geq 1$ (0 otherwise). Regions of solid colors correspond to regions of the parameter space producing a similar signal decay in this noiseless setting. The colored dots in the central parameter space correspond to the signal generated with the corresponding colored distribution (ground truth is red). As was the case with our previous two-cylinder example, we have a wide area of the parameter space generating roughly indistinguishable signals. The four distributions pictured on the sides all produce essentially identical signals for a wide range of distribution shapes.

2.5 Effective MR diameter

We have shown in the previous section (sec. 2.4) that it seems unfeasible to fit even the simplest distributions. Therefore, we might resort to fitting a single "effective" diameter. When fitting a single parameter over a quantity following a distribution, it is natural that this fitted value will take the form of a central tendency measure of that distribution (a "weighted average").

In the case of MR axon diameters, there are two main effects providing the

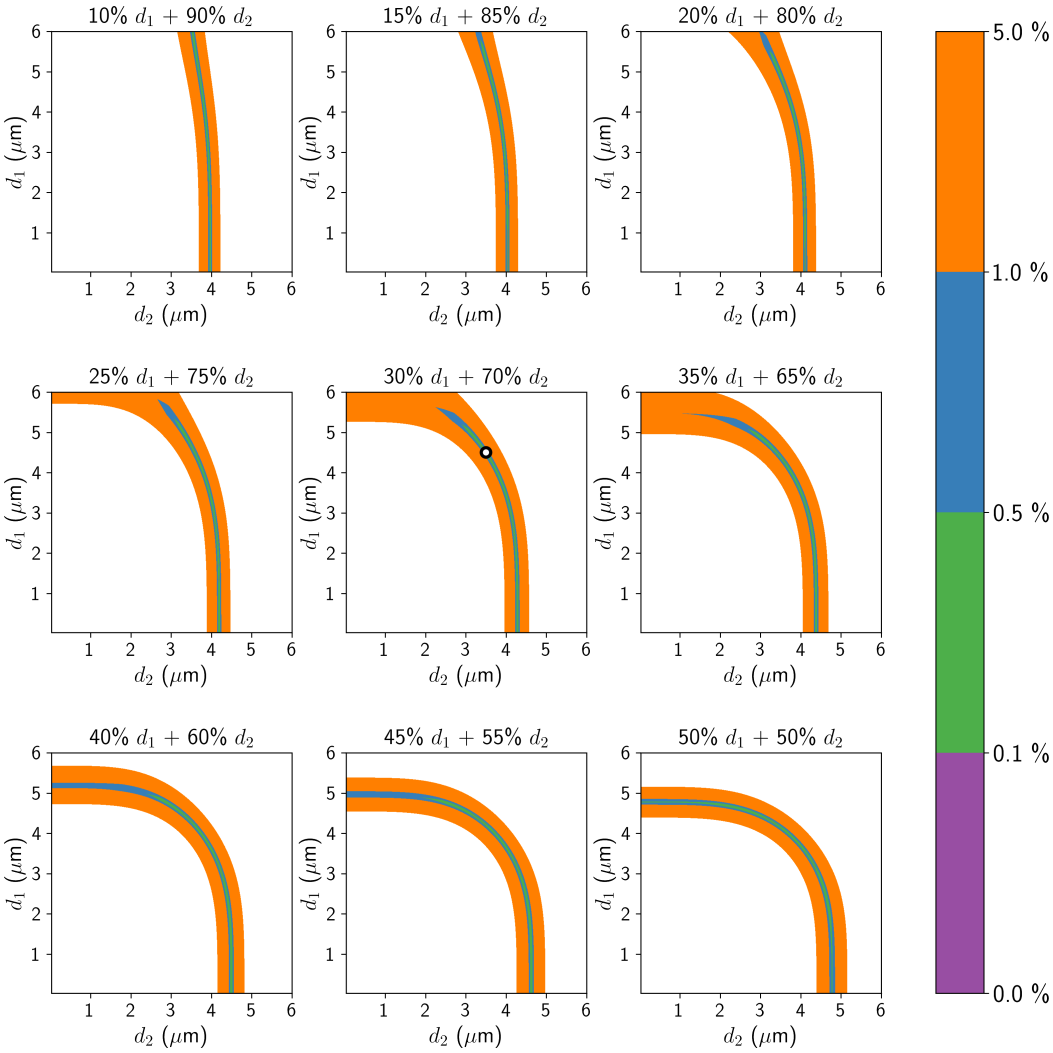

Figure 2: Example of the unresolvability of distribution fitting. The ground truth signal was generated from a combination of 2 parallel cylinders; 30% **signal fraction** with diameter $d_1 = 4.5$ μm and 70% $d_2 = 3.5$ μm (shown as white dot in the center plot) with in-vivo diffusivity ($D_0 = 2$ μm^2/ms) and a Connectom-like acquisition with three gradient pulse durations (G = 300 mT/m, $\Delta = 50$ ms, $\delta = [30, 40, 50]$ ms). The parameters were selected so that the smallest diameter was comfortably above "typical" diameter limit for $\delta = 30$ (compared to the limit for SNR = 30, this experiment is noiseless). The 9 subplots represent all combinations of diameters between 0.1 and 6 μm, sliced uniformly at signal fractions between 10% and 50%. The blue "path" correspond to parameter combinations yielding a signal less than 1% **signal decay** different than the noiseless ground truth. It forms a surface spanning most of the 3D parameter space, rendering any distribution fitting impossible for non-absurd SNR. Section A.4 showcase the same experiment for diameters closer to human axons.

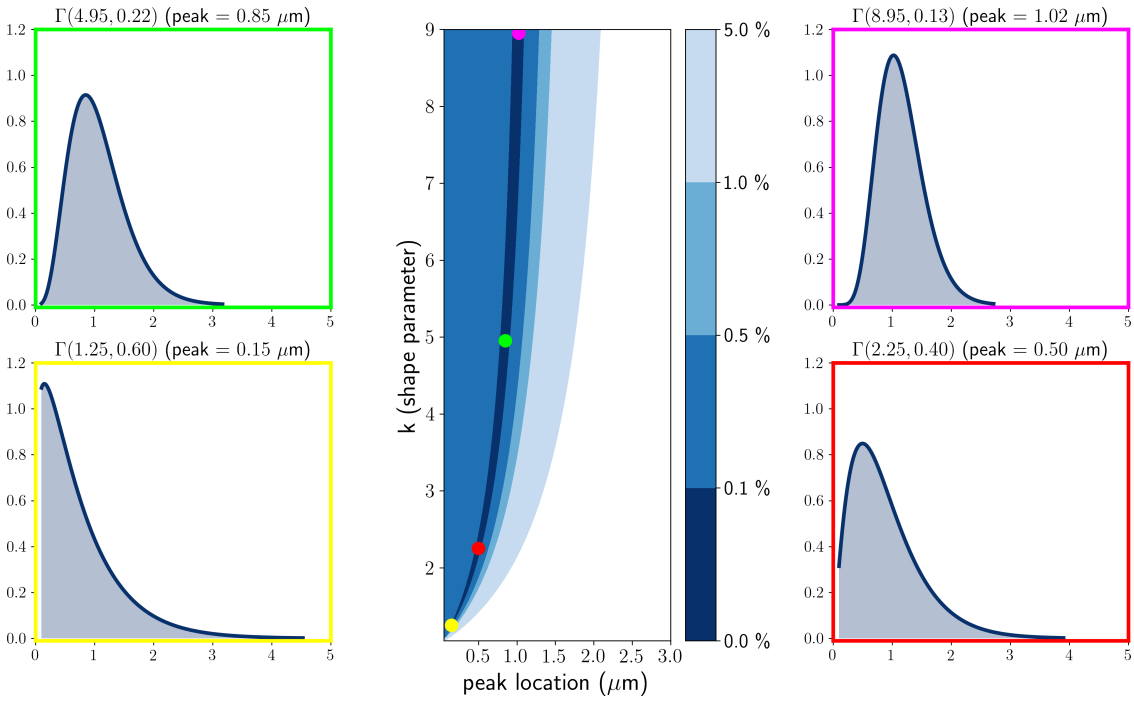

Figure 3: Example of the unresolvability of distribution fitting. The ground truth signal was generated using a gamma distribution of diameter count (shown as red dot in the center plot) with in-vivo diffusivity ($D_0 = 2\ \mu\text{m}^2/\text{ms}$) and a Connectom-like acquisition with three different gradient pulse durations (G = 300 mT/m, Δ = 50 ms, δ = [30, 40, 50] ms). The center plot represents all combinations of shape and peak location characterizing different gamma distributions. The dark blue "path" corresponds to parameter combinations yielding a signal less than 0.1% **signal decay** different than the noiseless ground truth. It forms a path spanning across most of the 2D parameter space, rendering distribution fitting unreliable for non-absurd SNR. The 4 side plots show examples of various gamma distributions from the center plot of wildly different shapes generating roughly indistinguishable signals.

"weighting". First, even though we are interested in the distribution of the *axon count*, the signal fractions are weighted by the *spin count*. Under the assumption of uniform intra-axonal proton density, T_2, same length cylinder for each diameter and no exchange, this manifest itself as a cross section area weighting, proportional to the 2nd power of the diameter. Secondly, the MR signal is sensitive to the 4th power of the diameter (as seen in eq. 4), adding up an extra heavy tail-weighting effect. Putting it all together, we can define the effective MR axon diameter d_{eff} over an arbitrary count distribution of density $P(d)$ as a function of its moments (eq. 2) [8, 38].

$$d_{\text{eff}} = \sqrt[4]{\frac{\langle d^6 \rangle}{\langle d^2 \rangle}} \tag{2}$$

where $\langle d^n \rangle = \int_{d'} P(d') d'^n \, \mathrm{d}d'$ is the nth moment of the distribution of density $P(d)$ (See sec A.2 for a simple proof-of-concept derivation). Fig. 4 shows a high match between the effective axon diameter computed from fitting a single diameter over the signal simulated from the distribution (d_{fit} in red) and the effective axon diameter derived from direct computation using the moments of the distribution (d_{eff} in green) for an example of a human axon diameter distribution from the left and right uncinate/inferior occipitofrontal fascicle taken from [24]. Preliminary post-mortem results [38] indicated a good correspondence between d_{eff} estimated from microscopy and from dMRI in a rat brain using a complex imaging strategy which properly suppresses non-intra-axonal signals and effects from axon orientations and dispersion.

Evidence points toward d_{eff} from eq. 2 being an accurate description of the "averaging" process of a typical dMRI sequence over a distribution of axons in the presence of no other signal [34, 38]. However, it is important to keep in mind the limitations of d_{eff} as a metric. By the nature of dMRI, it is extremely weighted toward the tail of the distribution as shown in fig. 4. The two distributions are fairly similar in terms of mean and peak location. However, the distribution of the left hemisphere (top plot) comprises an additional $\sim 2.5\%$ of large axons, effectively doubling the d_{eff} compared to the distribution of the right hemisphere (bottom plot). This difference in the distribution might be due to the small sample size inherent to histology. Such big axon outliers should not be an issue for dMRI where d_{eff} is estimated from voxels containing millions of axons. However, when comparing the d_{eff} metric between different ROIs or subjects, it becomes impossible to distinguish between situations such as a small global shift toward larger axons or a thicker distribution tail. This is expected when summarizing a complex distribution with only a single metric, instead of the two or three degrees of freedom it requires [33]. Nevertheless, the interpretability of d_{eff} is additionally impaired by the heavy tail weighting of its calculation, compared to other single metric distribution summary, such as the mean Fig. 5 shows the same axonal diameter distribution taken

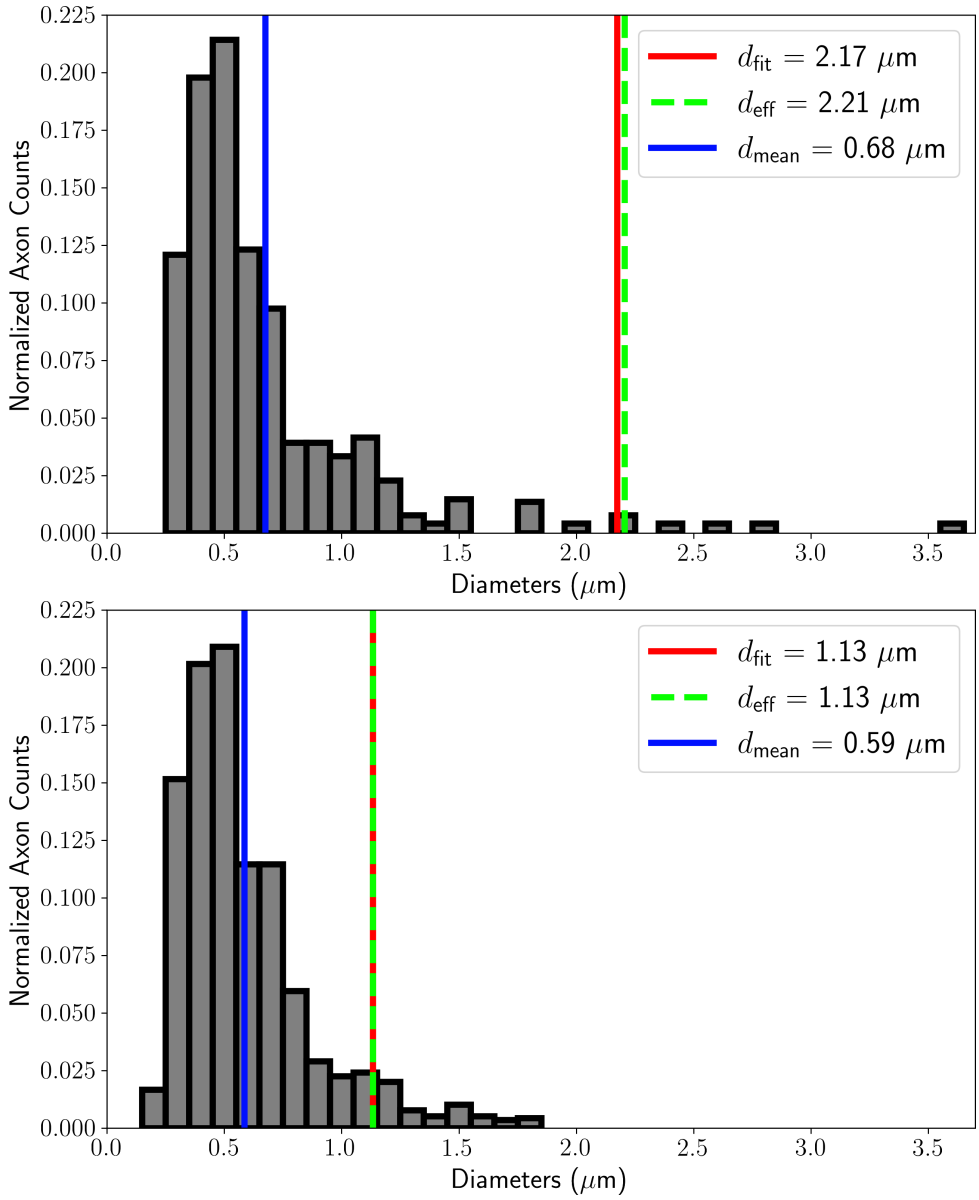

Figure 4: Human axon diameter normalized count distribution taken from Liewald et al.[24] (fig. 9, human brain 1, left and right hemisphere shown as top and bottom respectively). The peak diameter for both distributions is 0.5 μm while the mean diameter d_{mean} is around 0.6 μm. The bottom distribution maxes out below 2 μm while the top distribution has a few extra axons in the 2-4 μm range ($\sim$ 2.5% of axons by count). This small tail difference heavily affects the effective diameter d_{eff} (eq. 2) (doubles it in this case). The fitted MR diameter d_{fit} corresponds nicely with d_{eff} estimated from the moments of the distribution.

from Liewald et al.[24] overlapped with densities of multiple families of distributions (gamma, normal, uniform, exponential) with parameters tailored to produce the same theoretical d_{eff}. The goal is to clearly highlight the large (infinite) number of strikingly different distribution shapes that can produce the same d_{eff}. The interpretation of d_{eff} in its current state will require very strong hypothesis on the type of distributions or differences that can exist, which is not available in general.

3 Discussion and conclusion

The goal of this work is to showcase the sensitivity limits and the unresolvability of MR axon diameter models from PGSE diffusion weighted sequences. Those sensitivity limitations affect all dMRI diameter estimation models or techniques, as they are built upon cylindrical compartments and use the same restricted diffusion formulas. In section 2.2 and 2.3, we have shown how simple computations using realistic in-vivo parameters even with high-end Connectom MR gradient systems generate only very small signal decay with extremely limited sensitivity to relevant axonal diameters. Even the more favorable combination of post-mortem tissue and ultra-strong preclinical gradients does not result in sufficient signal decay to measure realistic axon diameters using diffusion MRI. The problem can be reframed statistically by comparing the signal decay to the noise level with a Z-test and defining a diameter limit. Computing $d_{\min}$ results in values that are very big compared to relevant axon diameters in the human brain. The effect of this limit was shown with an explicit simulation in fig. 1. In section 2.4, we have shown that fitting a distribution of diameters to the signal results in a multitude of widely different solutions even in the simplest settings. Finally, in section 2.5, we have shown how a distribution of diameters projects itself onto a single fitted diameter. While estimating d_{eff} from data seems feasible using advanced hardware and sequences [38], it remains a low dimensional and strongly tail-weighted projection of the distribution, making it ambiguous and insufficient for useful comparison.

We want to emphasize that every result in this work was computed utilizing idealized simulations that were arranged such that any presented limits correspond to a bound on the actual limit on real data. Notably, it has been shown [14, 27] that PGSE is the most sensitive sequence to axon diameters under the assumption of parallel cylinders and perpendicular diffusion gradient. In general, simulations are seen as insufficient to prove or disprove the effectiveness of a method. However, we work on data generated to perfectly follow the general assumptions of the fitted models and we observe that the model fits are still insensitive to the relevant underlying tissue properties. This implies serious doubts that any outputs from those models [5, 2, 18, 10, 17] applied to real data represent the underlying axonal diameter information, no matter how visually appealing they might seem. We suspect

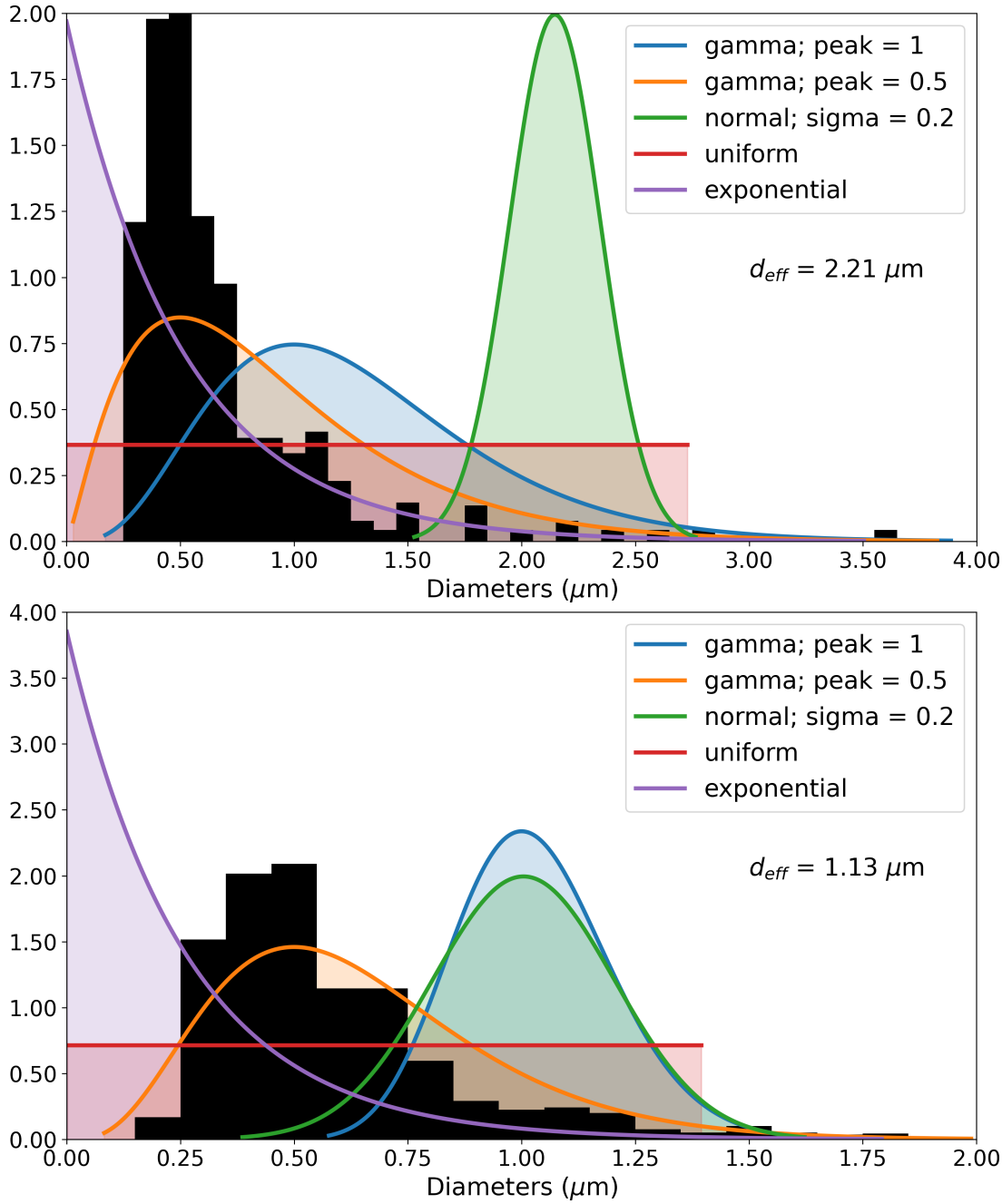

Figure 5: Different families of distributions tuned to produce the same d_{eff}. The target d_{eff} values were computed from the human axons diameter count distribution from Liewald et al.[24] (in black, the discrete counts were converted into a density for visualization). For both hemispheres, we used various families of distribution (restricted to be univariate) to show potential shape variance with identical d_{eff}.

that previous, apparently plausible results arise from some combination of effects outside of the model assumptions that are projecting themselves onto the model parameters in a complex way. Admittedly, some parameters can be reliably fitted to the data, such as the *effective diameter*, but as we have demonstrated such indices are insensitive to most relevant properties of the actual diameter distribution.

Hence, any claim of infeasibility of axon diameter measurement based on the employed simulations automatically translates to infeasibility of axon diameter measurements based on real data acquired with similar parameters. Our simulated data were generated (I) purely from intra-axonal signals and (II) perpendicular to the main orientation. In a multi-compartment model where the extra-axonal signal has to be fitted, (III) there will be residual fitting errors from the extra-axonal compartment contaminating the already tiny intra-axonal signal decay, increasing the effective $d_{\min}$. For example, a typical extra-axonal tensor compartment in the WM with a perpendicular diffusivity of 0.3 μm^2/ms produces a signal decay of $(1 - \exp(-4.28 * 0.3)) \approx 72.3\%$ for acquisitions parameters $\delta = \Delta = 10$ ms and $G = 300$ mT/m. If only 1% of this signal decay (i.e. 0.723% total signal decay) is instead erroneously considered as restricted compartment decay fitted with $D_0 = 2$ μm^2/ms, it would be equivalent to a cylinder with diameter 2.25 μm

In the simulations, we considered that the typical white matter SNR from an MR acquisition using Connectom gradients was driven only by intra-axonal compartment. (IV) However, in reality the intra-axonal volume fraction comprises less than 50% of the total volume in dense parallel fiber regions such as the corpus callosum and less in deep white matter [9]. This discrepancy (at least) halves the measured intra-axonal signal decay, thereby additionally increasing the effective $d_{\min}$. (V) Moreover, uncertainties in the estimation of the fiber orientation will additionally bias the apparent diameter because the restricted diffusion model will be fitted to the elongated elliptical cross-section. (VI) Unaccounted orientation dispersion for multi-compartment models will make estimation essentially impossible as shown in Nilsson et al.[27]. Considering all those sources of bias, it is clear that the already small signal decay caused by the restricted diffusion inside axons is essentially unattainable with such multi-compartment models.

An important message from eq. 1 and tables 1-2 are the scaling powers of the parameters. They are such that the sensitivity problem cannot be fixed using more powerful gradient systems. Even extreme cases such as going from in-vivo Connectom-like (G = 300 mT/m) acquisitions at normal SNR, to post-mortem measurements with ultra-strong preclinical gradients (G = 1500 mT/m) and 5 times better SNR (25 averages) only decreases the $d_{\min}$ from 2.56 to 0.58 μm (around 4.4 times better). This new value is barely enough to be sensitive to the peak of the diameter distribution in the best case. If we consider all the idealized assumptions from the diameter limit formula, it is likely not sufficient.

There are many misconceptions in the literature about the difficulty of going

from single diameter fitting to multiple diameters or a distribution. The "intuition" that errors in the fitted distribution will be normally distributed around the true solution fails spectacularly, even in the absolute simplest case of a signal from two axonal compartments with big diameters and no source of possible confounds as seen in fig. 2 and in section A.4. A commonly seen argument is to limit the distribution fit at some $d_{\min}$ best case value and claim that the resulting distribution must be valid because we are sensitive to these bigger diameters. Ignoring the value of $d_{\min}$, we can focus on what it fundamentally attempts to do, putting a limit on the minimal signal decay that can be statistically seen above the noise. To highlight this previous point, fig. 2 shows where configurations such as (35% 4.95 μm + 65% 2.9 μm), (30% 4.5 μm + 70% 3.5 μm), (100% 3.85 μm) and (45% 0.1 μm + 55% 4.6 μm) produced signal with [0.1, 0.5]% signal decay difference. Such a small decay requires SNR $\in [329, 1645]$ at optimal in-vivo Connectom-like settings, which correspond to a $d_{\min} \in [0.96, 1.44]$ μm, showing the disconnection between the limits of distribution fitting and direct $d_{\min}$ computation.

With the complexity of real axonal diameter distributions and the apparent impossibility of reliably fitting a distribution, working with the effective diameter d_{eff} seems to be the most promising avenue. When using an advanced acquisitions strategy to negate the non-intra-axonal signal [38], we can accurately estimate d_{eff}. However, d_{eff} is not a well "behaved" metric for comparisons between subjects or different brain areas. Before we can do such an analysis, we would potentially need to develop a new non-Stejskal-Tanner diffusion sequence producing a slightly different weighting of the distribution to allow some shape disentangling. In its current state, the metric d_{eff} cannot differentiate fundamentally different situations such as a small diameter increase of all axons versus a large diameter increase from a small proportion of the axon population.

An interesting topic we did not mention so far is the time-dependence of the extra-axonal space diffusion [28, 8, 23, 31]. Previous attempts to model axonal diameters assumed that all the time-dependent diffusivity portions of the signal were due to intra-axonal restricted diffusion. Recent work has highlighted a mechanism by which the extra-axonal space can also produce signals with time-dependent diffusivity. Indeed, the spacing of the restricting barrier in the extra-axonal compartment tends to be larger than typical axon diameters at relevant time-scales. This has the effect of producing a larger signal decay than the intra-axonal restricted compartment for a given acquisition scheme and to produce a time-dependent diffusivity when varying Δ. We briefly show in section A.3 how this extra-axonal time-dependence could contribute to the axon diameter overestimation seen in literature.

An apparent oversimplification throughout this work concerns how SNR and number of samples are chosen. For example, in fig. 1, our 1D approach is equivalent to generating the signal for a single gradient direction perpendicular to the cylin-

der. Similarly, we chose three directions for fig. 2 *i.e.* equal to the number of free parameters. If one has a real sample containing only identical parallel cylinders, the knowledge of the orientation would not be present and hundreds of directions would be sampled across multiple values of δ and G. It is hard to define a single value representing the SNR gain going from one data point with perfect alignment and with maximal sensitivity to hundreds of data points with varying sensitivity, extra parameters to fit and etc. If we take instead 100 repetitions of the optimal measurement and ignore the unknown orientations, we get an upper bound of $\sqrt{100} = 10$ times better SNR which corresponds to a $\sqrt[4]{10} = 1.78$ times smaller d_{min}. A more realistic upper bound is to include the estimation of the direction as two extra free parameters and frame the data as $\frac{100}{3}$ repetitions of three optimal measurements; $\sqrt{\frac{100}{3}} \approx 5.77$ times better SNR which corresponds to a $\sqrt[4]{5.77} \approx 1.52$ times smaller d_{min}. This view becomes increasingly complex as we add more parameters and start taking into account how different measurements have non-equal sensitivity to each of the estimated parameter. Since there is a 8th root scaling of d_{min} versus additional averaging (functional form of diameter versus signal decay is 4th power and SNR versus averages is 2nd power in the best case), we feel that results on minimal number of data points are sufficiently relevant.

The conduction velocity of myelinated axons is strongly impacted by axon diameter [39, 13]. The results presented in this work seem to indicate that direct conduction velocity estimations from dMRI are unfeasible, since they rely entirely on axon diameter results suffering from the sensitivity limitations. The overestimation of small to medium axon diameters translate into the overestimation of conduction velocities [20, 22]. Similarly, any error in compartment separation or main orientation estimation is likely to cause large bias in the diameter estimation. Under the most simplistic model, the conduction velocity of myelinated axons is a linear function of the outer axon diameter (*i.e.* including myelin sheath) [21], as well as the inner axon diameter under equal g-ratio. Under this relation, the conduction velocity estimated using d_{eff} is heavily weighted toward the conduction velocity of the largest axons, following the same pattern as Eq. 2 ($(\langle v^6 \rangle / \langle v^2 \rangle)^{1/4}$ for conduction velocities v) and suffering from all the same limitations.

In summary, our results show that the MR-based assessment of axonal diameters is methodologically infeasible. Our simulations under ideal conditions demonstrate that diffusion-weighted MRI with current and foreseeable future hardware is not capable of performing axonal diameter measurements in biologically relevant dimensions. The inability to measure axonal diameters is not a matter of the biophysical model choice but rather stems from the missing contrast of the intra-axonal tissue fraction. Under realistic, less ideal measurement conditions, the feasibility of such measurements is even further reduced. We show that frequently shown "known" variations of axonal diameter across structures such as the corpus callosum might also be explained with time dependent diffusion of the extra-axonal tissue frac-

tion. Therefore, previous measurements and model fitting results rather represent a characterization of the extra-axonal space than a measure or representation of the axonal diameter. Our manuscript further investigates recent descriptions of axonal diameters using a projection on an "effective diameter". Our simulations show this representation can be strongly affected by small changes in the distribution tail and does not allow to draw any unambiguous conclusions about the actual distribution of diameters.

Given the immense methodological difficulties of MR axonal diameter measurements, we suggest to include the time dependence of extra-axonal diffusion in the quantitative description of the microstructure of white matter in future studies. In connection with an independent measure of tissue myelination, this time dependency may provide an indirect approach to estimate the outer axonal diameter. Multidimensional dMRI measurements [36] may help to describe the extra-axonal space due to a reduced degeneracy of associated microstructural models. This may open a doorway to a quantitative study of brain microstructure using diffusion MRI.

Acknowledgement

MP is supported by a scholarship (PDF-502732-2017) from the Natural Sciences and Engineering Research Council of Canada (NSERC). MP and CE are supported by the Priority Program 2041 (SPP 2041) "Computational Connectomics" of the German Research Foundation (DFG).

A Appendices

A.1 Insensitivity of Δ to axon diameter

There is some misunderstanding in the literature concerning the impact of varying Δ to probe axon diameter. Intuitively, the dMRI signal is created by the dephasing of spins due to their displacement. For Δ to play a role in the measured restricted signal, we need to be in a short enough δ time regime. In the long time regime, by the end of the gradient application, most spins have interacted strongly with the axonal wall and their positions are mostly de-correlated from their initial position; the maximal signal decay has been reached and changing the gradient spacing Δ will not change anything. In the range of relevant parameter values (see sec. 2.1), it is simple to numerically show this phenomenon. Fig. 7 shows the signal decay computed from eq. 3 for all physically plausible (Δ, δ) pairs in $\Delta \in [10, 50]$ ms and $\delta \in [10, 50]$ ms for various axon diameters for an in-vivo Connectom-like settings. The respective signal decay depends strongly on the diameters, however, there is no perceptible difference for different Δ at the same δ. The same results can be achieved by Monte-Carlo spin diffusion simulation (see Fig. 6).

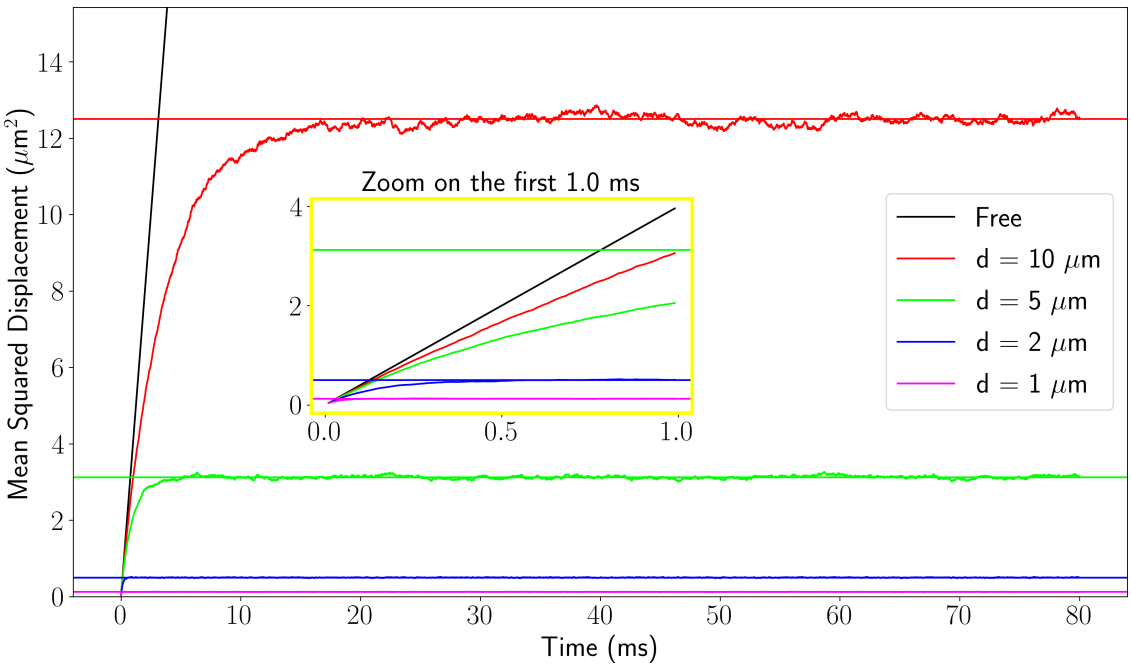

Figure 6: Mean squared displacement (MSD) for one direction from 2D Monte-Carlo simulation for free diffusion and restricted diffusion inside circles of different radius using $D_0 = 2\ \mu\text{m}^2/\text{ms}$. The horizontal lines show the long time limit MSD for each diameter. The center plot is a zoom on the first millisecond where we see that even the relatively large 2 μm diameter circle reaches long time regime quicker than any sufficiently strong diffusion gradient can be applied ($\delta_{\text{min}} \geq 5$ ms).

Another way to demonstrate this result is to derive the rough form of the signal equation from spin dephasing. We have applied gradient g and pulse width δ. In the long time regime, we have $\delta \gg t_c$, t_c being the characteristic correlation time of the cylinder ($t_c \sim d^2/D_0$). We will first calculate the phase ϕ_1 accumulated by spins within a time window of t_c (where the Gaussian phase approximation applies [26]) and then compute the total phase ϕ accumulated as a sum of $N \sim \delta/t_c$ uncorrelated contributions. Within one short step, phase is accumulated linearly proportional to the applied gradient and spin displacement, $\phi_1 \sim g d t_c$. We now compute the signal using $\ln(S) \sim -\phi^2 \sim -\phi_1^2 \delta/t_c = -g^2 d^2 t_c \delta = -\frac{g^2 d^4 \delta}{D_0}$. The recovered equation form corresponds to the Neuman long-time limit up to a constant and is independent of Δ and the initial position (it implicitly vanished by considering a displacement of d for a time-step of t_c in ϕ_1).

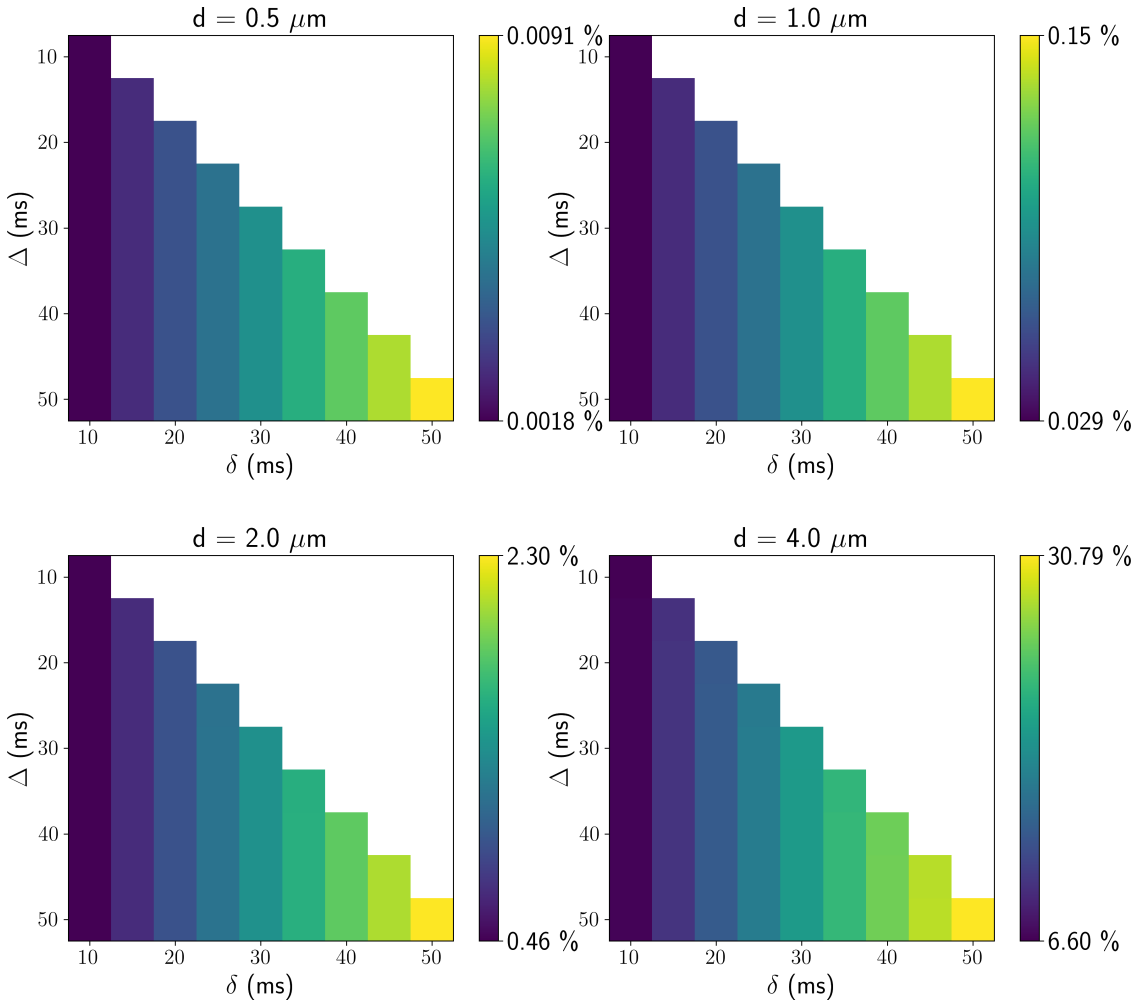

Figure 7: Noiseless MR signals from eq 3 for various (Δ, δ) and axon diameter (d). The signals were simulated for $G = 300$ mT/m and $D_0 = 2.0$ μm^2/ms. We note that different Δ (y-axis) doesn't modify the signal in any appreciable way.

The Van Gelderen formula (Eq. 3) for restricted diffusion is derived using the Gaussian phase approximation (GPA), which assume a Gaussian distribution for

spin dephasing. The GPA is guaranteed to hold for very short and very long diffusion time [26] and is empirically accurate for simple shapes of boundaries such as circles. It is a fast converging series, typically reaching precision errors smaller than 10^-7 with only 10 terms. Often, instead of truncating Eq. 3, we use the Neuman long-time limit formula (Eq. refeq:neuman1) for simplicity. It is typically numerically indistinguishable from Eq. 3 for relevant parameter range. The Neuman formula is defined assuming $\Delta = \delta$, but it can nonetheless be used in the case of the relevant parameter range where the restricted diffusion signal is insensitive to Δ

$$ln(E) = -2\gamma^2 G^2 \sum_{m=1}^{\infty} \left[\frac{2D_0\alpha_m^2\delta - 2 + 2e^{-D_0\alpha_m^2\delta}}{D_0^2\alpha_m^6((\frac{d}{2})^2\alpha_m^2 - 1)} + \frac{2e^{-D_0\alpha_m^2\Delta} - e^{-D_0\alpha_m^2(\Delta-\delta)} - e^{-D_0\alpha_m^2(\Delta+\delta)}}{D_0^2\alpha_m^6((\frac{d}{2})^2\alpha_m^2 - 1)} \right] \tag{3}$$

where E is the normalized diffusion signal, γ is the proton gyromagnetic ratio, G is the diffusion gradient amplitude, D_0 is the unrestricted diffusivity in the cylinder, Δ is the diffusion gradient separation, δ is the diffusion gradient duration, d is the diameter of the cylinder, α_m is the mth root of the equation $J_1'\left(\alpha \cdot \frac{d}{2}\right) = 0$, where $J_1'(\cdot)$ is the derivative of the Bessel function of the first kind.

$$E = \exp\left(-\frac{7}{1536}\frac{\gamma^2 G^2}{D_0} d^4 \left(2\delta - \frac{99}{448}\frac{d^2}{D_0}\right)\right) \tag{4}$$

A.2 Effective diameter derivation

We give a simple derivation of the effective diameter. The normalized MR signal as a function of d with all other parameters fixed is

$$E(d) = \exp\left(-\frac{7}{768}\frac{\gamma^2 G^2 \delta}{D_0} d^4\right) \equiv \exp\left(Cd^4\right)$$

for some fixed constant C. We compute the volume fraction normalized signal E_P for diameter **counts** following a distribution of density $P(d)$. We use the approximation $E(d) \approx 1 + Cd^4$ from the truncated Taylor series of $exp(\cdot)$.

$$
\begin{aligned}
E_P &= \int_d P(d) \frac{(\pi d^2)}{\int_{d'} P(d')(\pi d'^2)\,\mathrm{d}d'} E(d)\,\mathrm{d}d \\
&= \frac{\int_d P(d)(\pi d^2) E(d)\,\mathrm{d}d}{\int_d P(d)(\pi d^2)\,\mathrm{d}d} \\
&= \frac{\int_d P(d)(\pi d^2)(1 + C d^4)\,\mathrm{d}d}{\int_d P(d)(\pi d^2)\,\mathrm{d}d} \\
&= \frac{\int_d P(d)(\pi d^2)\,\mathrm{d}d}{\int_d P(d)(\pi d^2)\,\mathrm{d}d} + \frac{\int_d P(d)(\pi d^2) C d^4\,\mathrm{d}d}{\int_d P(d)(\pi d^2)\,\mathrm{d}d} \\
&= 1 + C \cdot \frac{\int_d P(d) d^6\,\mathrm{d}d}{\int_d P(d) d^2\,\mathrm{d}d} \\
&= 1 + C \cdot \frac{\langle d^6 \rangle}{\langle d^2 \rangle} \\
&= 1 + C \cdot \left(\sqrt[4]{\frac{\langle d^6 \rangle}{\langle d^2 \rangle}} \right)^4 \\
&= E\left(\sqrt[4]{\frac{\langle d^6 \rangle}{\langle d^2 \rangle}} \right) = E(d_{\text{eff}})
\end{aligned}
$$

■

A.3 Extra-axonal time-dependent diffusivity

It has been shown that the extra-axonal compartment can exhibit time-dependent diffusivity [28, 8, 23]. It arises from the disorder created by the irregular packing of axons of varying diameters. The "disorder strength" is characterized by the parameter A and has been empirically estimated in [8] to be $A \approx 0.2(l_c^{\perp})^2$ where $l_c^{\perp}$ is the fiber packing correlation length at which diffusion is restricted in extra-axonal space. Two models of perpendicular diffusivity as function of (Δ, δ) are described in [23]; $D_{\perp}^{\text{intra}}(\Delta, \delta)$ assuming that all the time dependence in the diffusivity arises from intra-axonal space, $D_{\perp}^{\text{extra}}(\Delta, \delta)$ assuming that all the time dependence in the diffusivity arises from the extra-axonal space.

$$
D_{\perp}^{\text{intra}}(\Delta, \delta) \simeq f_{ex} D_{\infty}^{ex} + \frac{c}{\delta(\Delta - \delta/3)}, \quad c = \frac{7}{768} \frac{f_{in} d_{\text{eff}}^4}{D_0} \tag{5}
$$

$$
D_{\perp}^{\text{extra}}(\Delta, \delta) \simeq f_{ex} D_{\infty}^{ex} + c' \frac{\ln(\Delta/\delta) + \frac{3}{2}}{\Delta - \delta/3}, \quad c' = f_{ex} A \tag{6}
$$

with extra-axonal volume fraction f_{ex}, intra-axonal volume fraction $f_{in} = 1 - f_{ex}$, long time ($\Delta \to \infty$) extra-axonal diffusivity D_{∞}^{ex}, bulk diffusivity D_0 and disorder strength parameter A.

Evidence on a few subjects suggest that the extra-axonal time-dependence dominates the intra-axonal time-dependence [31, 23]. This was shown by fitting both eq. 5 and 6 to data acquired with fixed $\delta = 20$ ms and multiple $\Delta \in [26, 100]$ ms

to comparable goodness-of-fit. The fitted parameters were then used to predict the signal values of a second acquisition using $\Delta = 75$ ms and multiple $\delta \in [4, 45]$ ms, where the extra-axonal model obtained good predictions and the intra-axonal model failed. Since most axon diameter estimation methods assume static values for the extra-axonal diffusivity, if the time-dependence in the signal is dominated by extra-axonal effects, the estimated diameters will be large and mostly unrelated to the effective diameter d_{eff}. To showcase this effect, we equated eq. 5 and 6 ($D_{\perp}^{\text{intra}}(\Delta, \delta) = D_{\perp}^{\text{extra}}(\Delta, \delta)$) and isolated d_{eff}. We used the typical value of $D_0 = 2$ μm^2/ms and fixed $D_{\infty}^{ex} = 0.5$ μm^2/ms (fitted values in [23] inside $[0.38, 0.6]$ μm^2/ms). We use $f_{ex} \in [0.25, 0.75]$ and $A \in [0.25, 2]$, giving us $f_{ex}A \in [0.0625, 1.5]$ compared to the reported values in [23] inside $[0.24, 0.56]$. We generated the *"fake"* d_{eff} for all physically plausible combinations of $\Delta \in [5, 100]$ ms and $\delta \in [5, 50]$ ms. We observe effective diameter between 2 μm and 9.5 μm, with most diameters above 6 μm in the configurations ($f_{ex} = 0.5$ and $A = [0.5, 1]$) closest to results from [23].

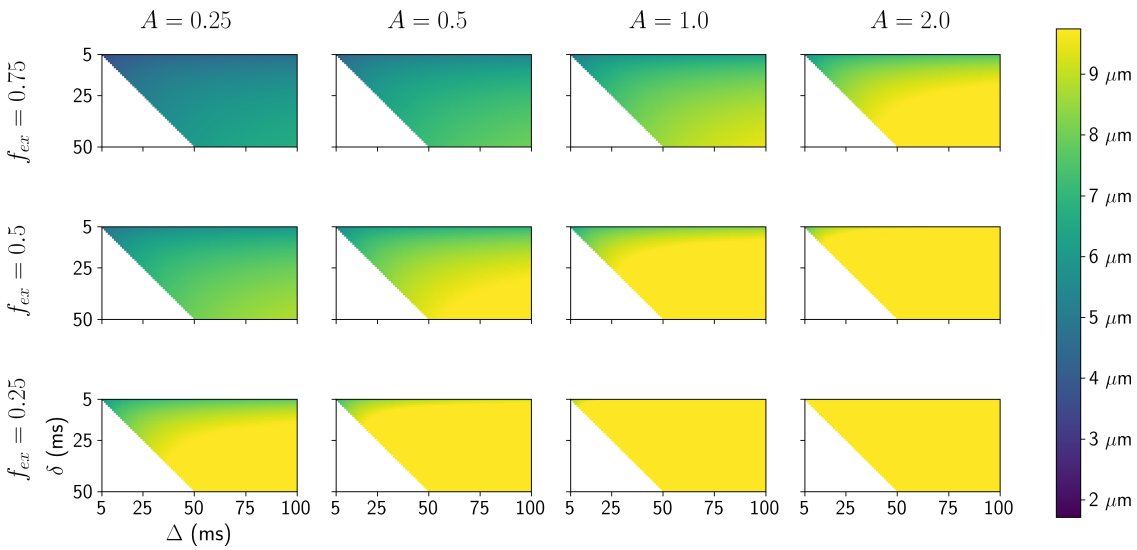

Figure 8: Signal generated using the extra-axonal time-dependence formula eq 6 and effective diameters fitted using eq 5.

The well-known "small-big-small diameter pattern" observed in the corpus callosum with histology and "reproduced" with big overestimation by axon diameter estimation methods ([2, 25, 16, 10]) can potentially be explained by this presented effect [31]. A brain area with a higher mean diameter is likely to also have an increased $l_c^{\perp}$ for random circle packing; if the diameter distribution is uniformly shifted up, the packing keeps the same relative efficiency and the individual inter space grows, alternatively, if a few more big axons are present, it increases the diameter heterogeneity and the packing efficiency tend to go down, creating more extra-axonal space. In any case, $f_{ex}A$ increases and the *"fake"* d_{eff} follows in the setting of fig. 8. However, the extra-axonal model parameters still contain some information about the *outer* diameter distribution, but it is complexly tangled with

axon packing.

A.4 Two-diameter distributions

We show more examples of fitting a two-diameter model with smaller, more realistic diameters. In fig. 2, we used a combination of enormous diameters (**signal fraction**, 30% $d_1 = 4.5$ μm and 70% $d_2 = 3.5$ μm) to highlight the effect of having a distribution over the lack of sensitivity of the realistic state-of-the-art acquisition scheme. We now show results for (30% $d_1 = 3.5$ μm and 70% $d_2 = 2.5$ μm) and (30% $d_1 = 2.5$ μm and 70% $d_2 = 1.5$ μm), where the ambiguity over the diameters is amplified for the same sampling scheme.

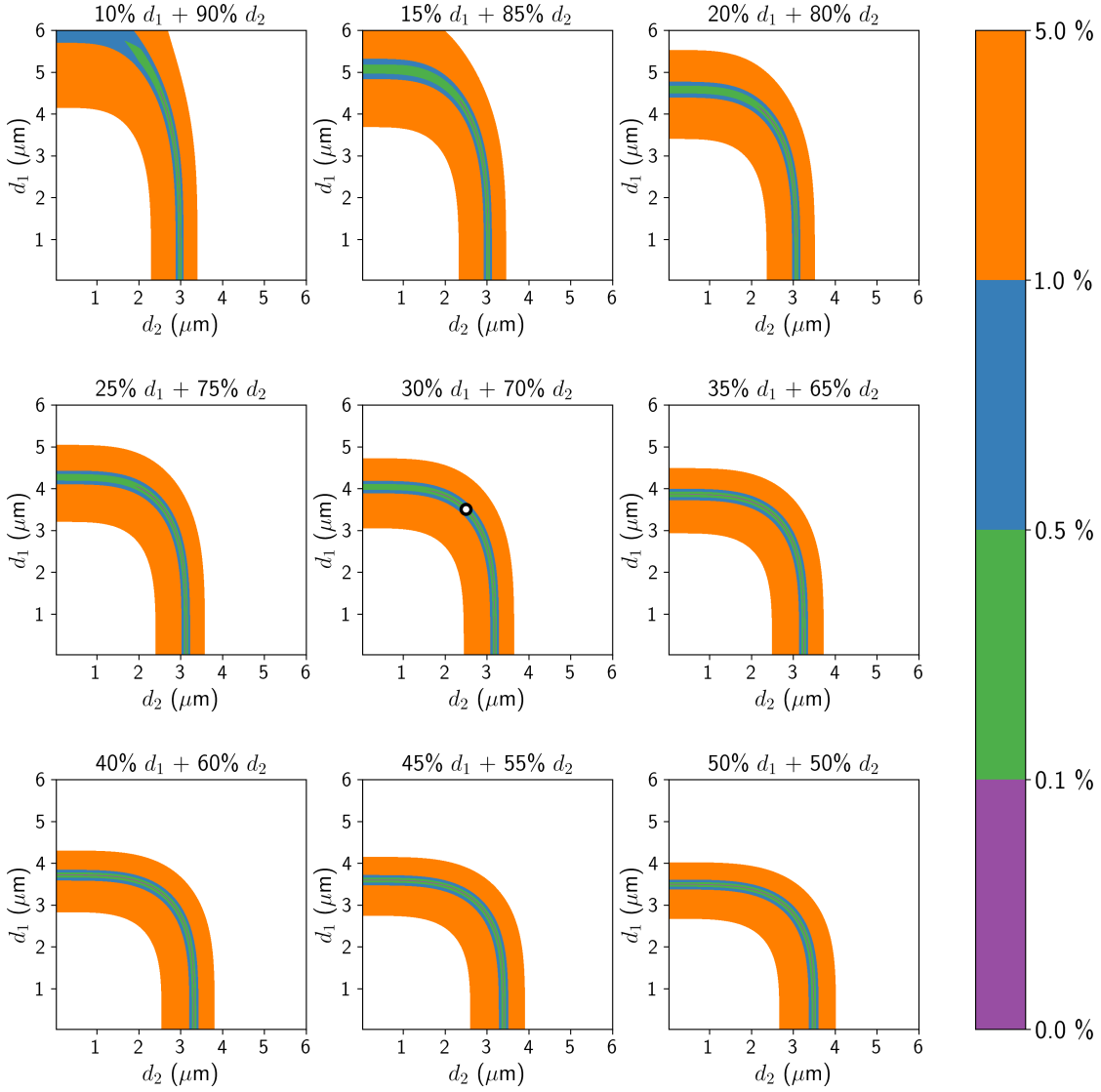

Figure 9: Example of the unresolvability of distribution fitting. The ground truth signal was generated from a combination of 2 parallel cylinders; 30% **signal fraction** with diameter $d_1 = 3.5$ μm and 70% $d_2 = 2.5$ μm (shown as white dot in the center plot) with in-vivo diffusivity ($D_0 = 2$ μm^2/ms) and a Connectom-like acquisition with three different gradient pulse durations (G = 300 mT/m, $\Delta = 50$ ms, $\delta = [30, 40, 50]$ ms). The parameters were selected so that the smallest diameter was comfortably above "typical" diameter limit for $\delta = 30$ (compared to the limit for SNR = 30, this experiment is noiseless). The 9 subplots represent all combinations of diameters between 0.1 and 6 μm, sliced uniformly at signal fractions between 10% and 50%. The blue "path" correspond to parameter combinations yielding a signal less than 1% **signal decay** different than the noiseless ground truth. It forms a surface spanning most of the 3D parameter space, rendering any distribution fitting impossible for non-absurd SNR.

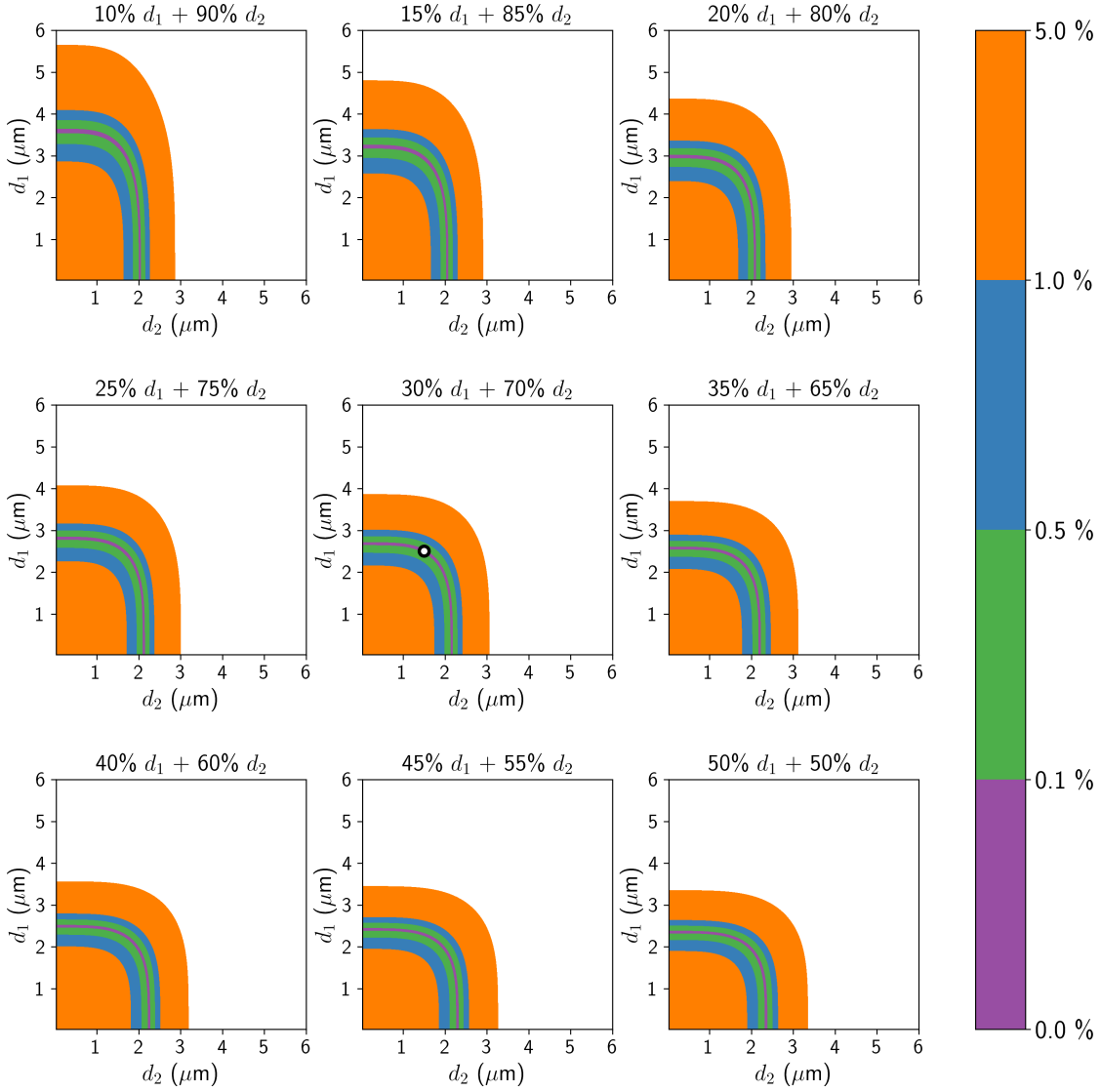

Figure 10: Example of the unresolvability of distribution fitting. The ground truth signal was generated from a combination of 2 parallel cylinders; 30% **signal fraction** with diameter $d_1 = 2.5$ μm and 70% $d_2 = 1.5$ μm (shown as white dot in the center plot) with in-vivo diffusivity ($D_0 = 2$ μm^2/ms) and a Connectom-like acquisition with three different gradient pulse durations (G = 300 mT/m, $\Delta = 50$ ms, $\delta = [30, 40, 50]$ ms). The parameters were selected so that the smallest diameter was comfortably above "typical" diameter limit for $\delta = 30$ (compared to the limit for SNR = 30, this experiment is noiseless). The 9 subplots represent all combinations of diameters between 0.1 and 6 μm, sliced uniformly at signal fractions between 10% and 50%. The blue "path" correspond to parameter combinations yielding a signal less than 1% **signal decay** different than the noiseless ground truth. It forms a surface spanning most of the 3D parameter space, rendering any distribution fitting impossible for non-absurd SNR.

Bibliography

References

[1] F. Aboitiz, A. B. Scheibel, R. S. Fisher, and E. Zaidel. Fiber composition of the human corpus callosum. *Brain Research*, 598(1):143–153, 1992.

[2] D. C. Alexander, P. L. Hubbard, M. G. Hall, E. A. Moore, M. Ptito, G. J. M. Parker, and T. B. Dyrby. Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage*, 52(4):1374–89, 2010.

[3] M. Ashtarayeh, T. Streubel, J. Periquito, A. Pohlmann, T. Niendorf, E. Kirilina, M. Morawski, C. Jäger, S. Geyer, M. Barakovic, A. Daducci, and S. Mohammadi. Axon diameter estimation in fixed human optic chiasm using diffusion weighted MR microscopy and microstructure-informed tractography. *ISMRM*, 2019.

[4] Y. Assaf and P. Basser. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage*, 27(1):48–58, 2005.

[5] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P. J. Basser. AxCaliber: A method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine*, 59(6):1347–1354, 2008.

[6] M. Barakovic, G. Girard, D. P. R. Romascano, J. R. Patino Lopez, M. Descoteaux, G. Innocenti, D. K. Jones, J.-P. Thiran, and A. Daducci. Assessing feasibility and reproducibility of a bundle-specific framework on in vivo axon diameter estimates at 300mT/m. *ISMRM*, 2018.

[7] M. Barakovic, D. P. R. Romascano, G. Girard, M. Descoteaux, J.-P. Thiran, and A. Daducci. In-vivo bundle-specific axon diameter distributions estimation across the corpus callosum. *ISMRM*, 2017.

[8] L. M. Burcaw, E. Fieremans, and D. S. Novikov. Mesoscopic structure of neuronal tracts from time-dependent diffusion. *NeuroImage*, 114:18–37, 2015.

[9] S. Coelho, J. M. Pozo, M. Costantini, J. R. Highley, M. Mozumder, J. E. Simpson, P. G. Ince, and A. F. Frangi. Local volume fraction distributions of axons, astrocytes, and myelin in deep subcortical white matter. *NeuroImage*, 179:275–287, 2018.

[10] A. Daducci, E. J. Canales-Rodríguez, H. Zhang, T. B. Dyrby, D. C. Alexander, and J.-P. Thiran. Accelerated microstructure imaging via convex optimization (AMICO) from diffusion mri data. *NeuroImage*, 105:32 – 44, 2015.

[11] A. Daducci, A. Dal Palù, A. Lemkaddem, and J. P. Thiran. Commit: Convex optimization modeling for microstructure informed tractography. *IEEE Transactions on Medical Imaging*, 34(1):246–257, 2015.

[12] B. Dhital, M. Reisert, E. Kellner, and V. G. Kiselev. Intra-axonal diffusivity in brain white matter. *NeuroImage*, 189:543–550, 2019.

[13] M. Drakesmith, R. Harms, S. U. Rudrapatna, G. D. Parker, C. J. Evans, and D. K. Jones. Estimating axon conduction velocity in vivo from microstructural mri. *NeuroImage*, 203:116186, 2019.

[14] I. Drobnjak, H. Zhang, A. Ianuş, E. Kaden, and D. C. Alexander. Pgse, ogse, and sensitivity to axon diameter in diffusion mri: Insight from a simulation study. *Magnetic Resonance in Medicine*, 75(2):688–700, 2016.

[15] T. B. Dyrby, W. F. Baaré, D. C. Alexander, J. Jelsing, E. Garde, and L. V. Søgaard. An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. *Human Brain Mapping*, 32(4):544–563, 2011.

[16] T. B. Dyrby, L. V. Sogaard, M. G. Hall, M. Ptito, and D. C. Alexander. Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI. *Magnetic Resonance in Medicine*, 70(3):711–721, 2013.

[17] Q. Fan, A. Nummenmaa, B. Wichtmann, T. Witzel, C. Mekkaoui, W. Schneider, L. L. Wald, and S. Y. Huang. Validation of diffusion MRI estimates of compartment size and volume fraction in a biomimetic brain phantom using a human MRI scanner with 300 mT/m maximum gradient strength. *Neuroimage*, 182:469–478, 11 2018.

[18] 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. *Sci Rep*, 6:38927, 12 2016.

[19] U. Grenander. On empirical spectral analysis of stochastic processes. *Arkiv för Matematik*, 1(6):503–531, 1952.

[20] A. Horowitz, D. Barazany, I. Tavor, M. Bernstein, G. Yovel, and Y. Assaf. In vivo correlation between axon diameter and conduction velocity in the human brain. *Brain Struct Funct*, 220(3):1777–1788, 2015.

[21] J. B. Hursh. Conduction velocity and diameter of nerve fibers. *American Journal of Physiology-Legacy Content*, 127(1):131–139, 1939.

[22] G. M. Innocenti, R. Caminiti, and F. Aboitiz. Comments on the paper by Horowitz et al. (2014). *Brain Struct Funct*, 220(3):1789–1790, May 2015.

[23] H.-H. Lee, E. Fieremans, and D. S. Novikov. What dominates the time dependence of diffusion transverse to axons: Intra- or extra-axonal water? *NeuroImage*, 182:500–510, 2018.

[24] D. Liewald, R. Miller, N. Logothetis, H.-J. Wagner, and A. Schüz. Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque. *Biological Cybernetics*, 108(5):541–557, 2014.

[25] J. A. McNab, B. L. Edlow, T. Witzel, S. Y. Huang, H. Bhat, K. Heberlein, T. Feiweier, K. Liu, B. Keil, J. Cohen-Adad, M. D. Tisdall, R. D. Folkerth, H. C. Kinney, and L. L. Wald. The Human Connectome Project and beyond: Initial applications of 300mT/m gradients. *NeuroImage*, 80:234–245, 2013.

[26] C. H. Neuman. Spin echo of spins diffusing in a bounded medium. *Journal of Chemical Physics*, 60(11):4508–4511, 1974.

[27] M. Nilsson, S. Lasič, I. Drobnjak, D. Topgaard, and C.-F. Westin. Resolution limit of cylinder diameter estimation by diffusion MRI: The impact of gradient waveform and orientation dispersion. *NMR in Biomedicine*, 30(7):e3711, 2017.

[28] D. S. Novikov, E. Fieremans, S. N. Jespersen, and V. G. Kiselev. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR in Biomedicine*, 32(4):e3998, 2019.

[29] A. Roebroeck, K. L. Miller, and M. Aggarwal. Ex vivo diffusion MRI of the human brain: Technical challenges and recent advances. *NMR in Biomedicine*, 32(4):e3941, 2019.

[30] W. A. H. Rushton. A theory of the effects of fibre size in medullated nerve. *The Journal of Physiology*, 115(1):101–122, 1951.

[31] S. D. Santis, D. K. Jones, and A. Roebroeck. Including diffusion time dependence in the extra-axonal space improves in vivo estimates of axonal diameter and density in human white matter. *NeuroImage*, 130:91–103, 2016.

[32] H. Schmidt and T. R. Knösche. Action potential propagation and synchronisation in myelinated axons. *PLOS Computational Biology*, 15(10):1–33, 2019.

[33] F. Sepehrband, D. C. Alexander, K. A. Clark, N. D. Kurniawan, Z. Yang, and D. C. Reutens. Parametric Probability Distribution Functions for Axon Diameters of Corpus Callosum. *Front Neuroanat*, 10:59, 2016.

[34] F. Sepehrband, D. C. Alexander, N. D. Kurniawan, D. C. Reutens, and Z. Yang. Towards higher sensitivity and stability of axon diameter estimation with diffusion-weighted MRI. *NMR Biomed*, 29(3):293–308, Mar 2016.

[35] E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 42(1):288–292, 1965.

[36] D. Topgaard. Multidimensional diffusion MRI. *Journal of Magnetic Resonance*, 275:98–113, 2017.

[37] P. van Gelderen, D. Despres, P. C. M. van Zijl, and C. T. W. Moonen. Evaluation of restricted diffusion in cylinders. phosphocreatine in rabbit leg muscle. *Journal of Magnetic Resonance - Series B*, 103(3):255–260, 3 1994.

[38] J. Veraart, D. Nunes, U. Rudrapatna, E. Fieremans, D. K. Jones, D. S. Novikov, and N. Shemesh. Noninvasive quantification of axon radii using diffusion MRI. *eLife*, 9:e49855, 2020.

[39] S. G. Waxman. Determinants of conduction velocity in myelinated nerve fibers. *Muscle Nerve*, 3(2):141–150, 1980.

[40] H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, and D. C. Alexander. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61(4):1000–1016, 2012.