"All explanations are theory laden" – David Deutsch in The Beginning of Infinity

Q1.1.1 code cells 6 – 16. We plot the mean diffusivity map in Fig. 1 and FA map with directionally encoded color map in Fig. 2. We interpret them in their captions.

Q1.1.2 code cells 17-24. Before fitting, we'd expect to have a RESNORM of $N * \sigma^2 = 108 * (200)^2 = 4.3e6$ because if we have a perfect fitting model the RESNORM is equal to the sum of the noise error squared (i.e. variance). Using the given starting parameters (`x = [3.3e3,0.001,0.45,1,1]`), we obtain `x: [3.5e3, -5.1e-6, 1.2e2, 8.9e-1, 1.6]` with a RESNORM that's 5 times higher what we expect, as well as the estimated diffusivity f is wrongly negative because it should be positive as it represents the volume fraction of the isotropic diffusion compartment. We visualize this unexpected bad-quality fit in Fig. 3 left. In Fig. 3 right we observe a better fit using a starting point given in the lectures `x2=[4200, 4e-4, 0.25, 0, 0]` with RESNORM `5.9e6`. We obtain sensible parameters `x=[4.3e+3,1.1e-3,3.6e-1,-9.8e-1,5.8e-1]`, except for a negative θ which is wrong because it has to be positive since it represents the orientation of the fibers in the brain tissue.

Q1.1.3 code cells 25-32. We apply the following transformations $x = \left[S(0,0)^2, diff^2, logit(f), logit\left(\frac{\theta}{\pi}\right), logit\left(\frac{\phi}{2\pi}\right)\right]$ to ensure parameters respect the following constraints for biophysical plausibility: $S(0,0) > 0$ because it's the non-DW signal, $diff > 0$ because it refers to the magnitude of water diffusion from high to low concentration, $f \in [0,1]$ because it's the volume fraction of water undergoing anisotropic diffusion, and $\theta \in (0,\pi), \varphi \in (0,2\pi)$, since they're angles defining the orientation of diffusion of water along axon fibers. Numerically, θ and φ can be any value on the real domain but constraining them increases the chances of finding a local minimum faster (as later experiments show), as well as ensures that estimated parameters θ, ϕ have a unique mapping in the specified domain, making it easy to check if we have found the same fitted parameters in later questions. We use the starting point of `x2=[4200, 4e-4, 0.25, 1, 1]`, we find `x = [4257,0.00114,0.36,2.2,0.58]` which results in the same RESNORM as in Q1.1.2 probably because this is the global minimum for this voxel. Different to Q1.1.2, now we have biophysical plausible parameters. The parameters are also similar in values except θ. See Fig 4.

Q1.1.4 code cells 33-36. Because it is impossible to know a priori where the global minimum is, other than it exists for a SSD loss function, we repeat the model fitting procedure $N = 100$ iterations, each with a random starting point computed via adding some random Gaussian noise to our original starting point used in Q1.1.3 (`x2=[4200, 4e-4, 0.25, 1, 1]`), with each noise dimension using a standard deviation $\sigma = x2/5$. This is to explore realistic starting points, such as avoiding values biophysically implausible (e.g. $S(0,0) \ll 3000$). Through this approach applied to 6 random voxels, we obtain for one voxel `[85,55]` a smallest RESNORM of `3.2e6`, found 30% of the time. However, we reiterate this may not be the global minimum. Assuming the lowest RESNORM per voxel is the "global minimum", we write $P(finding\_global\_minimum) = p = 0.3$. We then model each run with a Bernoulli distribution of finding the global min or not, where we calculate the number of times we need to do a run to find the global min with a probability of 0.95 as $\frac{log(0.05)}{log(1-p)}$. For example, for this voxel we need 9 runs to have a 95% chance of finding the global min. Repeating this for a few other voxels we find a range of N runs required, the highest being 9, therefore across the whole image slice we will use 9 runs. Results from other voxels are inconsistent, e.g., for voxel [71, 70], I can reach the global minimum 98% of the time, highlighting the sensitivity of the ball-stick model. See Table 1 for results.

Q1.1.5 code cells 37-40: The estimated parameter maps from our constrained optimization routine, RESNORM map and fiber directions are plotted in Fig 5 and 6. In the RESNORM and diffusivity maps there are some white spots indicating a bad fit.

Q1.1.6: code cells 41-56. I try 4 experimental settings: informed/uninformed starting points with constrained/unconstrained) optimization using 6 random voxels where I measure the probability of finding the "global minimum". The informed starting points are computed by mapping the linearly fit DT model's parameters as follows: the eigenvectors of the DT represent the principal directions of diffusion, and the eigenvalues give the diffusivity along these directions. Given the principal eigenvector v, ball-stick parameters can be computed as follows: $f = FA, diff = \lambda_1; \theta = \arccos\left(\frac{v_z}{\sqrt{v_x^2+v_y^2+v_z^2}}\right); \phi = arctan2(v_y, v_z)$, where the volume fraction f is approximated with the fractional anisotropy as they both quantify degree of anisotropy. The $diff$ is approximated by the axial diffusivity which corresponds to the largest eigenvalue λ_1 of the principal eigenvector decomposed from the DT model's parameters. The spherical coordinates result from trigonometric transformations of the components of the principal eigenvector v. Running the experiments with 6 random voxels, we find **informed starting points with constrained optimization** maximizes the likelihood of finding the best parameters in the least, expected amount of runs per voxel ($N = 4$) with often $> 80\%$ to do so. In comparison, uninformed starting points in Q1.1.3-1.1.5, either constrained or not, need $N = 14$ runs per voxel to do so. We proceed in computing the parameter maps and fiber directions for the informed starting points with constrained optimization in Figures 7 & 8, where we note that it took ~25 mins compared to the ~30 mins for constrained, uninformed optimization in Q1.1.5. We next compute the derivative analytically (coupled with constrained optimization) for 6 random voxels and find that it greatly helps with finding the global minima almost all the time with the same expected amount of runs per voxel ($N = 4$), like informed optimization but with the addition of quicker computation times. The parameter map with informed starting point generated via this approach is in Figures 9 and 10. Other than the diffusivity map, all other scalar parameters as well as the fiber directions display a very bad fit denoted by the plethora of white spots. See Table 2 for results.

Q1.1.7 code cells 58-66. We replace the Gaussian SSD with the Rician negative log-likelihood and perform constrained optimization with uninformed starting points over 6 random voxels. We find all parameters estimated using the Rician likelihood are like the ones using the Gaussian counterpart. The exception is S0 which is orders of magnitude smaller than the Gaussian-estimated S0 (`~1e-13 vs ~4e3`). We next compute the parameter maps (~10 mins) and notice they are of worse quality, as well as we can't observe any fibers compared to the one estimated using Gaussian likelihood. We argue this could be due to wrong implementation of the Rician NLL since it's very difficult to adjust.

Q1.2.1 code cells 70-72**.** Classical bootstrap is non-parametric, and thus makes no assumption on the data distribution. We thus sample uniformly with replacement the voxel intensities. Such sampling yields 95% confidence level and 2-sigma range estimates for the fitted parameters. For voxel $[92, 65]$, the bootstraps are in Table 3, with bootstrap distributions of each parameter for 4 random voxels in Fig 13. In the figure we can see the distributions are clearly not normal because the 95%**.** We particularly note a heavily right-skewed, non-Normal distribution for f, and arguably this is because at these voxels, there is an abundance of axons pointing at different directions, thus there is a high-volume fraction of diffusion.

Q1.2.2 code cells 73-76. We implement the MCMC algorithm, where we note that we modified the perturbing noise to around 120 to get an acceptance rate of 23%, which is appropriate when we're

estimating the posterior $P(x|A)$ of ≥ 3 parameters of the ball-stick model. Comparing the MCMC with the bootstrap output at voxel [92, 65], we notice in the MCMC our 2σ ranges match with their respective 95% grey region better than the bootstrap, especially for f where in the MCMC we do observe a Gaussian-like distribution (Figure 14). MCMC estimates and plots are in Table 4. In Figure 15, we compare how parameters estimated from the bootstrap predicts differently to the ones estimated using MCMC. We note that perhaps due to random sampling, the bootstrap estimates only fit MRI signal data of values around 2500, while MCMC was able to look at the entire dataset to account for signal intensities of around 3000-4000.

Q1.2.3 79-82 We compare Laplace method (Figure 16), classical (non-parametric) bootstrap of Q1.2.1 and parametric bootstrap assuming Rician data distribution (Figure 17) for voxel [92, 65]. We notice that for parametric bootstrap with Rician assumption yields better estimates than non-parametric counterpart, particularly for $S0$ which is not stuck at around ~2500 as before (see Table 5). We also note that in terms of computational time, Laplace is intensive due to the Hessian's inverse bottleneck, but only needs to be computed once, while bootstrapping (parametric or not) requires iterative computations, and thus takes longer.

Q1.2.4 code cells 83-89**.** We draw inspiration from Harms & Roebreck (2018) and use MCMC by plotting the posterior mean of parameters mapped over a brain slice and use the corresponding std. deviation to visualize uncertainty in Figure 18 and plot fiber maps with uncertainty in Figure 19 where the plotted uncertainty is measured via propagation of the standard deviation when computing the fiber directions.

Q1.3.1 code cells 91-94. We choose a starting point of `x4=[4.2, 4e-4, 0.25, 0.9, 1]` which is the same as in Q1.1.3, except that S0 is scaled down since the new dataset contains measurements orders of magnitude different to before. We perform constrained optimization using the transformations as in Q1.1.3. Now we expect a RESNORM of $N * \sigma^2 = 5.8$., however witch each of the 6 voxels get about 3 to 4 times larger than this (see Table 6). The plotted predictions are in Fig. 20:

Q1.3.2 code cells 95-102**.** The benchmarked models' fit to the data can be observed in Figures 21, 22, 23: The best RESNORMS for Ball-stick, DT, Zeppelin & Stick (ZS) and ZS with Tortuosity are, respectively: 15.1 at voxel 1, 16.3 at voxel 6, 10.8 at voxel 1, and 12.9 at voxel 1.

Q1.3.3 code cells 103-105. AIC and BIC rankings remain consistent between one another (Table 7). The ZS consistently achieves the lowest scores, indicating it fits the data better with a reasonable model complexity w.r.t to the other candidates.

Q1.3.4 code cells 106-108**.** We reuse an implementation online of the Ball-2-Stick model that performs constrained optimization. We tried different initializations, but we get extremely poor fits, with lowest RESNORM at 87 for voxel 1. The results are plotted in Fig 24.

Q1.3.5 code cells 103-105**.** The Zeppelin & Stick model consistently achieves the highest performance compared to variants in all 6 voxels as per the AIC and BIC.

Q1.3.6 code cells 109-114. We plot the parameter maps estimated from the ZS and ZS-with Tortuosity models where we highlight in red regions in the S0 map where the AIC of the ZS is higher than the ZS with Tort. (Fig. 25). I argue that the Tortuosity component improves the fit to the data because it counts for the variability in the diffusion paths due to the brain's microstructural environment, such complex arrangement of axons. Despite being more complex than the ZS model, the likelihood is too high that obfuscates the influence of number of parameters.

References:

Harms, R., & Alard Roebroeck. (2018). Robust and Fast Markov Chain Monte Carlo Sampling of

Diffusion MRI Microstructure Models. *Frontiers in Neuroinformatics*, *12*. https://doi.org/10.3389/fninf.2018.00097