Non central chi estimation of multi-compartment models improves model selection by reducing overfitting

Aymeric Stamm[1], Benoit Scherrer[1], Stefano Baraldo[2], Olivier Commowick[3], and Simon Warfield[1]

[1]*Computational Radiology Laboratory, Boston Children's Hospital, Boston, MA, United States,* [2]*MOX, Politecnico di Milano, Milan, Italy,* [3]*VISAGES, INRIA, Rennes, France, Metropolitan*

PURPOSE – Diffusion images are known to be corrupted with a non-central chi (NCC)-distributed noise [1]. There has been a number of proposed image denoising methods that account for this particular noise distribution [1,2,3]. However, to the best of our knowledge, no study was performed to assess the influence of the noise model in the context of diffusion model estimation as was suggested in [4]. In particular, multi-compartment models [5] are an appealing class of models to describe the white matter microstructure but require the optimal number of compartments to be known a priori. Its estimation is no easy task since more complex models will always better fit the data, which is known as over-fitting. However, MCM estimation in the literature is performed assuming a Gaussian-distributed noise [5,6]. In this preliminary study, we aim at showing that using the appropriate NCC distribution for modelling the noise model reduces significantly the over-fitting, which could be helpful for unravelling model selection issues and obtaining better model parameter estimates.

METHODS – Noise in diffusion MRI is NCC-distributed as $f\left(S_j|A_j,S_0,\sigma^2\right)=\frac{S_0A_j}{\sigma^2}\left(\frac{S_j}{S_0A_j}\right)^L e^{-\frac{S_j^2+S_0^2A_j^2}{2\sigma^2}}I_{L-1}\left(\frac{S_jS_0A_j}{\sigma^2}\right)$, where L and σ^2 are respectively the effective number of coils and effective variance of each coil that depend on the coils' correlations, the acceleration factor of the acquisition and the signal reconstruction method [3,4], S_j is the reconstructed measured signal, S_0 is the unknown non-diffusion-weighted signal, A_j is the signal attenuation predicted by the model and I_{L-1} is the $(L-1)$-order modified Bessel function of the first kind. We set up an NCC estimation framework that can be used to fit any diffusion model that provides an analytic expression of the predicted signal attenuation A_j. Starting from N acquired diffusion images, we compute the NCC log-likelihood by replacing A_j with the appropriate expression and then we maximize it using a derivative-free optimizer. Researchers usually rely on the assumption of high SNR to resort to the Gaussian approximation of the noise. However, at high SNRs, the NCC log-likelihood reads $\log\mathcal{L}=\log\left((2\pi\sigma^2)^{-1/2}exp\left(-\sum_{j=1}^{N}\left(S_j-S_0A_j\right)^2/(2\sigma^2)\right)\right)-(2L-1)/2\sum_{j=1}^{N}\log\left(S_0A_j/S_j\right)$, which does not boil down to the Gaussian log-likelihood (blue term). Indeed, the red term penalizes any model that predicts a signal greater than the observed one, i.e., any model that predict physically implausible signals. For smaller SNRs, the penalty can even dominate the likelihood. Gaussian noise does not worry for these physical incoherences.

RESULTS – The objective of the experimental design was to demonstrate that accounting for the real nature of the noise in diffusion images helps in unravelling the model selection problem by reducing over-fitting. We used two cases provided by the Human Connectome Project [7]. For each case, the data consisted in 270 DW images with three different b-values at 1000, 2000 and 3000 s/mm2 and 18 non-DW images, at a spatial resolution of $1.25 \times 1.25 \times 1.25$ mm^3. We estimated the same ball-and-stick models with increased number of fascicle compartments using both Gaussian and NCC (with L=32 degrees of freedom) noise models and performed a model selection based on Akaike information criterion for each noise model to output maps of the optimal number of fascicles voxelwise. We proposed a qualitative assessment of the resulting maps and we reported, for each noise model, the proportion of voxels where 0, 1, 2 and 3 fascicles were detected.

DISCUSSION – We compared the resulting maps of optimal number of fascicles obtained by means of AIC on both subjects. Figure 1 shows a coronal view of these maps for subject #1 (a,b) and #2 (c,d) assuming Gaussian (a,c) and NCC (b,d) noise. These examples qualitatively show that assuming Gaussian noise instead of NCC noise induces a lot of over-fitting that favors the 3-fascicle model. As a result, 1-fascicle regions are much better delineated using NCC noise. Specifically, using the exact same fitting procedure with only an improved noise model, the proportion of 0,1,2-fascicle voxels (models with low complexity) is increased and the proportion of 3-fascicle voxels is decreased as summarized in Table 1.

# Fascicles	0	1	2	3
Gaussian (%)	9	4	27.5	59.5
NCC (%)	18.5	7.5	32	42

Table 1: Averaged proportion of 0,1,2,3-fascicle voxels across subjects

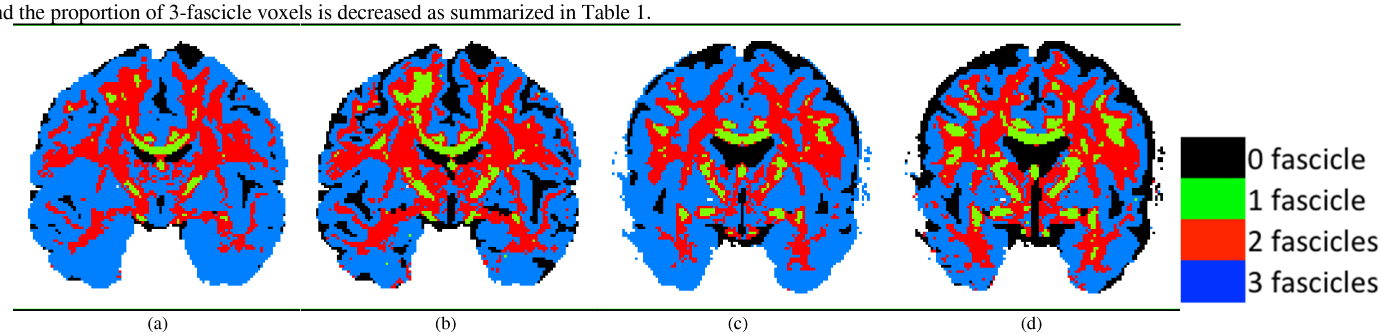

Figure 1: Coronal view of the maps of optimal number of fascicles for subject #1 (a,b) and #2 (c,d): assuming Gaussian noise (a,c) and NCC noise (b,d).

CONCLUSION – We set up an NCC framework for the estimation of multi-compartment models and performed a preliminary study to understand whether it is important to account for the real NCC nature of the noise or if the usual Gaussian approximation yields equivalent results. In this work, we specifically compared the maps of optimal number of fascicles obtained assuming Gaussian noise or NCC noise with degrees of freedom equal to the number of coils, using the same ball-and-stick model estimated with the same maximum likelihood procedure and the same optimization algorithm. Results show that accounting for the right distribution of the noise is crucial for model selection, since assuming a Gaussian noise yields to many more voxels with over-fitting. Future works will aim at (i) evaluating the consequences on the estimated parameters and (ii) estimating the spatially varying number of coils to use the correct NCC noise distribution in each voxel [3,4].

REFERENCES – [1] Aja-Fernandez et al., Noise estimation in single- and multiple-coil magnetic resonance data based on statistical models, Magn. Reson. Med., 2009, [2] Rajan et al., Nonlocal maximum likelihood estimation method for denoising multiple-coil magnetic resonance images, Magn. Reson. Imaging, 2012, [3] Aja-Fernandez et al., Effective noise estimation and filtering from correlated multiple-coil MR data, Magn. Reson. Imaging, 2012 [4] Aja-Fernandez et al., Statistical noise analysis in GRAPPA using a parametrized noncentral Chi approximation model, Magn. Reson. Med., 2011, [5] Panagiotaki et al., Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison, NeuroImage, 2012, [6] Scherrer et al., Parametric representation of multiple white matter fascicles from cube and sphere diffusion MRI, PloS one, 2012, [7] Sotiropoulos et al., Effects of image reconstruction on fibre orientation mapping from multi-channel diffusion MRI: reducing the noise floor using SENSE, Magn. Reson. Med., 2013, [HCP Data] Data were provided by the HCP, WU-Minn Consortium funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.