

DISCUSSION OF “FIBER DIRECTION ESTIMATION IN DIFFUSION MRI”

BY ARMIN SCHWARTZMAN

University of California, San Diego

1. Introduction. The paper by Wong, Lee, Paul and Peng offers a telling example of experimentalist vs. statistical thinking. While the experimentalists in the imaging world proposed to solve the problem of identifiability of the multi-tensor model by making more measurements at higher spatial resolution and at multiple magnetic gradient strengths, Wong et al. asked if estimation was still possible in the usual experimental setting and offered a different solution. By choosing the right parametrization, they have shown that, while the full multi-tensor model cannot be estimated, the diffusion directions can. If the goal of the analysis is tractography, then this is sufficient.

The paper is comprehensive, taking the analysis all the way from the measured diffusion directions to fiber tracking. Estimation of the diffusion directions is followed by spatial smoothing to improve accuracy. Clustering is used so that smoothing is only applied within fibers where the angle between neighboring voxels is small. This helps the fiber tracking algorithm better survive the difficult fiber crossing regions.

Because the objects of the analysis are diffusion directions, it is of interest to connect the work by Wong et al. to the existing body of knowledge of directional statistics. In the rest of this comment, I explore how directional statistics may shed some additional light on the problem.

2. The multi-tensor model. A key idea in the paper is that, while the full multi-tensor model is unidentifiable, the tensor eigenvectors are not. The identifiability problem is hard to spot at first because the multi-tensor model [equation (1) in Wong et al.] looks very much like a mixture of Gaussians. Gaussian mixture models are generally identifiable unless there is degeneracy in the parameters. The model is deceiving, however, because the argument vector \mathbf{u} is constrained to have unit norm. In fact, if the displacement of water molecules is modeled as a Gaussian mixture distribution, then the signal model in equation (1) is proportional to the Fourier transform of this distribution, restricted to a sphere of constant radius in frequency space [Mori (2007)]. The manipulation by Scherrer and Warfield (2010) (described by Wong et al. in Section 3.1) indicates that it is possible to modify the diffusion tensors in the original Gaussian mixture together with the mixture

weights in such a way that the values of the Fourier transform on the sphere remain constant.

Because the restriction of the Fourier transform does not discard directional information, diffusion directions can still be estimated. To reduce the number of parameters, Wong et al. choose the axially symmetric tensor model in their equation (2). To understand this model, it is illustrative to consider the molecule displacement distribution that gives rise to it. This is a Gaussian mixture where each component j has covariance matrix proportional to $\mathbf{D}_j = \alpha_j \mathbf{m}_j \mathbf{m}_j^\top + \xi_j \mathbf{I}$. Using the binomial inverse theorem for matrices to invert \mathbf{D}_j , the exponent in the corresponding Gaussian density of the displacement $\mathbf{x} \in \mathbb{R}^3$ is proportional to

$$-\mathbf{x}^\top \mathbf{D}_j^{-1} \mathbf{x} = -\mathbf{x}^\top \frac{1}{\xi_j} \left(\mathbf{I} - \frac{\alpha_j \mathbf{m}_j \mathbf{m}_j^\top}{\alpha_j + \xi_j} \right) \mathbf{x} = \frac{1}{\xi_j} \left(-\mathbf{x}^\top \mathbf{x} + \frac{\alpha_j (\mathbf{x}^\top \mathbf{m}_j)^2}{\alpha_j + \xi_j} \right).$$

Making the change of variable $\mathbf{x} = r \mathbf{u}$ where $r = |\mathbf{x}|$ and $|\mathbf{u}| = 1$, it follows that, given r , \mathbf{u} has a distribution on the unit sphere with positive exponent

$$\kappa_j (\mathbf{u}^\top \mathbf{m}_j)^2$$

for some $\kappa_j > 0$ (proportional to r^2). In other words, the direction of the molecular displacement for each tensor follows a Watson distribution with mean \mathbf{m}_j and concentration parameter κ_j [Watson (1965); Schwartzman et al. (2008)]. The distribution of diffusion directions in the axially symmetric multi-tensor model is thus a Watson mixture on the sphere.

3. Averaging of diffusion directions. Viewing the distribution of diffusion directions as a Watson mixture helps guide the following steps, for example, that of smoothing directions among neighboring voxels. While the Watson mixture model above describes the distribution of diffusion directions *within* a voxel, it may also describe the distribution of diffusion directions *between* neighboring voxels. After all, the division of the brain into voxels is an arbitrary artifact of the imaging technique and the anatomical structures in the brain span scales both smaller and larger than a voxel.

From this point of view, we may model diffusion directions belonging to the same fiber in neighboring voxels as draws from a single component of the Watson mixture. It is sensible to define the average direction as the maximum likelihood estimator of the mean vector in that component. This estimator, from the form of the Watson density and using Wong et al.'s notation in their equation (6), can be written as

$$\arg \max_{\mathbf{v} \in \mathcal{M}} \sum_{i=1}^T w_i (\hat{\mathbf{m}}_i^\top \mathbf{v})^2 = \arg \min_{\mathbf{v} \in \mathcal{M}} \sum_{i=1}^T w_i [1 - (\hat{\mathbf{m}}_i^\top \mathbf{v})^2].$$

The optimization problem above is similar to that in equation (6) of Wong et al. and is also a Karcher mean, but it uses the orthogonal projection distance $d^\perp(\mathbf{u}, \mathbf{v}) =$

$[1 - (\mathbf{u}^T \mathbf{v})^2]^{1/2}$ instead of the arc distance $d^*(\mathbf{u}, \mathbf{v}) = \arccos(|\mathbf{u}^T \mathbf{v}|)$ chosen by the authors.

Numerically, this may not have a large effect if the directional distribution is well concentrated around the mean. If $\theta = d^*(\mathbf{u}, \mathbf{v})$, then $d^\perp(\mathbf{u}, \mathbf{v}) = \sin \theta$, and the two are approximately the same if θ is small. Moreover, the arc distance has an appealing interpretation in terms of angle. However, the choice of distance has an effect computationally. While the minimization problem posed by Wong et al. requires an algorithmic solution based on geodesic coordinates and an elaborate accompanying theoretical justification, the quadratic minimization above based on the Watson distribution has a direct solution. In fact, the problem translates to

$$\arg \max_{\mathbf{v} \in \mathcal{M}} \mathbf{v}^T \left[\sum_{i=1}^T w_i (\hat{\mathbf{m}}_i \hat{\mathbf{m}}_i^T)^2 \right] \mathbf{v},$$

whose solution is simply the eigenvector corresponding to the largest eigenvalue of the weighted empirical covariance matrix in the large parentheses [Watson (1965); Schwartzman et al. (2008)]. This direct solution may help simplify and shorten computation time in Wong et al.'s Algorithm S1 (Step 6) and Algorithm S4 (Steps 7 and 10), which may be substantial over hundreds of thousands of voxels. It may also simplify Theorem 2.

4. Fiber tracking. Another advantage of the Watson average is that it comes with a measure of uncertainty, called angle dispersion in Schwartzman et al. (2008). This would help quantify the uncertainty in the tractography algorithm, which Wong et al. did not include, and even lead to a Kalman filter type of tractography that takes into account this uncertainty in the estimation. This may be an interesting direction for follow-up work.

ADDITIONAL REFERENCE

WATSON, G. S. (1965). Equatorial distributions on a sphere. *Biometrika* **52** 193–201. [MR0207115](#)

DIVISION OF BIostatISTICS
UNIVERSITY OF CALIFORNIA, SAN DIEGO
9500 GILMAN DR., MC0631
LA JOLLA, CALIFORNIA 92093-0631
USA
E-MAIL: armins@ucsd.edu