Free Water Elimination and Mapping from Diffusion MRI

Gabriel Samberg

September 2025

1 Introduction

In this paper a way of restoring and separating meaningful signal from noisy and distorted data in Diffusion MRI via the Beltrami framework is proposed. A significant drawback of DTI (Diffusion Tensor Imaging) of the brain is that in each voxel the mean water displacement is being represented. Hence, in voxels that represent partial brain tissue areas such as areas on the contour of the brain that are in direct contact with free water or areas around an edema in the brain, the isotropic diffusion of the free water volume dominates the signal, and hence the water diffusion of the brain tissue in that area is mixed with a high signal contributed by the free water compartment.
In this review, I will mainly focus on the mathematical framework proposed in the paper. I will present the mathematical approach to this problem and derive in detail the calculations omitted from the paper.

2 Theory and definitions

In this section, I will go over the mathematical framework of the paper and the different notations used throughout.

We denote by $\mathbf{D}$ the diffusion tensor. At each spatial location $(x, y, z) \in \Omega$, $\mathbf{D}[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ is a 3×3 Positive Definite matrix which represents the diffusion of water in the voxel located at (x, y, z) in the scan.

The set $\{q_1, \ldots, q_n\}$ is the set of applied gradient directions which was used in the process of acquiring the Diffusion MRI data.

$A_{tissue}(\cdot)$ and $A_{water}(\cdot)$ are both vectors in $\mathbb{R}^n$ where the i'th coordinate measures water displacement in the direction of q_i of tissue compartment and water compartment respectively. We have

$$A_{tissue}(\mathbf{D})_k = exp(-bq_k^T \mathbf{D} q_k), \qquad A_{water}(\mathbf{D})_k = exp(-bd)$$

Where d is the apparent diffusion coefficient (ADC) of free water and b is the diffusion weightings.

The bi-tensor model $A_{bi\_tensor}(\cdot, \cdot)$ defined as

$$A_{bi\_tensor}(\mathbf{D}, f) = fA_{tissue}(\mathbf{D}) + (1 - f)A_{water}$$

Gives us an indication on the tissue to water ratio (via the parameter f) and information on the anisotropic diffusion of water in the tissue compartment (via $A_{tissue}(\mathbf{D})$).

3 Overcoming the ill posedness of the Bi-tensor model by regularization

The objective is to find a modeled attenuation of a bi-tensor that will minimize the distance to the measured attenuation tensor $\hat{A}$. namely we are looking for

$$(\tilde{\mathbf{D}}, \tilde{f}) = argmin_{\mathbf{D},f} ||A_{bi\_tensor}(\mathbf{D}, f) - \tilde{A}||_2^2 \qquad (1)$$

The problem is that there is an infinite number of possible solutions. Notice that for every $f \in (0,1)$ we can take the vector

$$\tilde{A}_t = \frac{\tilde{A} - (1-f)A_{water}}{f}$$

And see that

$$||(f\tilde{A}_t + (1-f)A_{water}) - \tilde{A}||_2^2 = 0$$

Solving for $\tilde{A}_t$ requires solving a system of linear equations that determines a single unique tensor $\tilde{\mathbf{D}}_t$ and hence there are infinite number of couples $(\mathbf{D}, f)$ that minimize (1). In order to narrow down the possible optimal solutions to a set of more desirable ones, we apply additional constraints on the problem and demand a piece-wise smooth continuity between neighboring voxels. This restriction is based on the biological limitations that we expect an organic tissue to posses. Hence, we add the smoothness regularization term $\sqrt{|\gamma(\mathbf{D})|}$ to our functional that we desire to minimize. Here γ is the induced metric on the spatial domain from the spatial-feature space in which the scan in being embedded.

4 The mathematical approach via the Beltrami framework point of view

We view a Diffusion MRI scan as an embedded manifold in $\mathbb{R}^9$ by the isometry

$$X : (\Omega, \gamma_{\mu\nu}) \longrightarrow (M, h_{ij})$$

Where $\Omega \subset \mathbb{R}^3$ is the scan spatial domain, M is a surface embedded in the spatial-feature space (in our case the spatial-feature space is $\mathbb{R}^3 \times P_3$, where P_3 is the space of positive definite matrices), h is a metric on M and γ is the induced metric on Ω (which essentially makes X an isometry).
We then define a functional on the space of such embedded surfaces and minimize it using variational calculus. In our case, we choose to minimize the functional with respect to the feature coordinates of the embedding space, and hence we keep the spatial coordinates unchanged.
The functional proposed in the paper is defined as

$$S(\mathbf{D}, f) = \int_\Omega ||A_{bi\_tensor}(\mathbf{D}, f) - \tilde{A}||_2^2 + \alpha\sqrt{|\gamma(\mathbf{D})|}d\Omega$$

Where $\mathbf{D}$ is the diffusion tensor defined earlier and effectively is in the feature part of the embedded surface M,f is a scalar parameter which we also minimize with respect to, and $\sqrt{|\gamma(\mathbf{D})|}$ is the smoothness regularization term mentioned in the previous section.

5 Calculating The metric on the spatial-feature space

The paper uses The Iwasawa decomposition in order to parametrize the space of positive definite matrices. Notice that the Iwasawa decomposition states that every invertible matrix M can be written as

$$M = N^T AH$$

Where N is upper triangular matrix, A is diagonal with positive entries on the diagonal, and H is an orthogonal matrix.
Now, given a positive definite matrix $P \in P_n$ (P_n is the space of real Positive definite matrices of size nxn), we know that it admits Cholesky decomposition, and hence

$$P = LL^T$$

Using Iwasawa decomposition on L we get that

$$P = LL^T = (N^T AH)(N^T AH)^T = N^T AHH^T A^T N = N^T A^2 N$$

Hence, using Iwasama decomposition and the fact that positive definite matrices admit a Cholesky decomposition, we can use the following parametrization of the space of 3x3 positive definite matrices

$$F : \mathbb{R}^3_{>0} \times \mathbb{R}^3 \longrightarrow P_3$$

Given by

$$F(x_1, x_2, x_3, x_4, x_5, x_6) = Y = N^T A N = \begin{pmatrix} 1 & 0 & 0 \\ x_4 & 1 & 0 \\ x_5 & x_6 & 1 \end{pmatrix} \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix} \begin{pmatrix} 1 & x_4 & x_5 \\ 0 & 1 & x_6 \\ 0 & 0 & 1 \end{pmatrix}$$

In the paper, a natural Riemannian metric was chosen on P_3

$$ds^2 = Trace((Y^{-1}dY)^2), \qquad Y \in P_3$$

Now, we would want to derive the metric matrix given the parametrization with Iwasawa coordinates. Is this case

$$ds^2 = Trace((Y^{-1}dY)^2) = Trace(((N^T AN)^{-1} d(N^T AN))^2)$$

We are therefore interested in the coefficients of $dx_1^2, dx_2^2, \dots$ in ds^2.
We begin by using Leibnitz rule and noticing that

$$dY = d(N^T AN) = d(N)^T AN + N^T d(A)N + N^T Ad(N)$$

Hence, we have the following equality

$$Y^{-1}dY = N^{-1}A^{-1}N^{-T}(d(N)^T AN + N^T d(A)N + N^T Ad(N)) = N^{-1}A^{-1}N^{-T}d(N)^T AN + N^{-1}A^{-1}d(A)N + N^{-1}d(N)$$

We first calculate the basic components of the above equality

$$N^{-1} = \begin{pmatrix} 1 & -x_4 & x_4x_6 - x_5 \\ 0 & 1 & -x_6 \\ 0 & 0 & 1 \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} \frac{1}{x_1} & 0 & 0 \\ 0 & \frac{1}{x_2} & 0 \\ 0 & 0 & \frac{1}{x_3} \end{pmatrix}$$

$$dN = \begin{pmatrix} 0 & dx_4 & dx_5 \\ 0 & 0 & dx_6 \\ 0 & 0 & 0 \end{pmatrix}, \qquad dA = \begin{pmatrix} dx_1 & 0 & 0 \\ 0 & dx_2 & 0 \\ 0 & 0 & dx_3 \end{pmatrix}$$

Now, we will calculate $Y^{-1}dY$ in three parts by calculating separately $N^{-1}A^{-1}N^{-T}d(N)^T AN$, $N^{-1}A^{-1}d(A)N$, and $N^{-1}dN$.
Omitting here the long calculations, we arrive at the following

$$N^{-1}dN = \begin{pmatrix} 0 & dx_4 & dx_5 - x_4 dx_6 \\ 0 & 0 & dx_6 \\ 0 & 0 & 0 \end{pmatrix},$$

$$N^{-1}A^{-1}d(A)N = \begin{pmatrix} \frac{dx_1}{x_1} & x_4(\frac{dx_1}{x_1} - \frac{dx_2}{x_2}) & x_5(\frac{dx_1}{x_1} - \frac{dx_3}{x_3}) + x_4x_6(\frac{dx_3}{x_3} - \frac{dx_2}{x_2}) \\ 0 & \frac{dx_2}{x_2} & x_6(\frac{dx_2}{x_2} - \frac{dx_3}{x_3}) \\ 0 & 0 & \frac{dx_3}{x_3} \end{pmatrix},$$

$$N^{-1}A^{-1}N^{-T}(dN)^\top AN = \begin{pmatrix} \begin{gathered} \Big[-\frac{x_1x_4}{x_2} - \frac{x_1}{x_3}(x_4x_6^2 - x_5x_6)\Big] dx_4 \\ + \frac{x_1}{x_3}(x_4x_6 - x_5)\, dx_5 \end{gathered} & \begin{gathered} \Big[-\frac{x_1x_4^2}{x_2} - \frac{x_1}{x_3} x_4x_6(x_4x_6 - x_5)\Big] dx_4 \\ + \frac{x_1x_4}{x_3}(x_4x_6 - x_5)\, dx_5 + \frac{x_2}{x_3}(x_4x_6 - x_5)\, dx_6 \end{gathered} & \begin{gathered} \Big[-\frac{x_1x_4x_5}{x_2} - \frac{x_1}{x_3} x_5x_6(x_4x_6 - x_5)\Big] dx_4 \\ + \frac{x_1x_5}{x_3}(x_4x_6 - x_5)\, dx_5 + \frac{x_2x_6}{x_3}(x_4x_6 - x_5)\, dx_6 \end{gathered} \\ \Big[\frac{x_1}{x_2} + \frac{x_1x_6^2}{x_3}\Big] dx_4 - \frac{x_1x_6}{x_3}\, dx_5 & \begin{gathered} \Big[\frac{x_1x_4}{x_2} + \frac{x_1x_4x_6^2}{x_3}\Big] dx_4 \\ - \frac{x_1x_4x_6}{x_3}\, dx_5 - \frac{x_2x_6}{x_3}\, dx_6 \end{gathered} & \begin{gathered} \Big[\frac{x_1x_5}{x_2} + \frac{x_1x_5x_6^2}{x_3}\Big] dx_4 \\ - \frac{x_1x_5x_6}{x_3}\, dx_5 - \frac{x_2x_6^2}{x_3}\, dx_6 \end{gathered} \\ -\tfrac{x_1x_6}{x_3}\, dx_4 + \tfrac{x_1}{x_3}\, dx_5 & -\tfrac{x_1x_4x_6}{x_3}\, dx_4 + \tfrac{x_1x_4}{x_3}\, dx_5 + \tfrac{x_2}{x_3}\, dx_6 & -\tfrac{x_1x_5x_6}{x_3}\, dx_4 + \tfrac{x_1x_5}{x_3}\, dx_5 + \tfrac{x_2x_6}{x_3}\, dx_6 \end{pmatrix}.$$

From here, we calculate

$$Trace((Y^{-1}dY)^2) = Trace((N^{-1}A^{-1}N^{-T}d(N)^TAN + N^{-1}A^{-1}d(A)N + N^{-1}dN)^2)$$

The calculation is very long, and hence I omit it as well; nevertheless, after calculating, we arrive at the following final expression

$$ds^2 = Trace((Y^{-1}dY)^2) = \frac{dx_1^2}{x_1^2} + \frac{dx_2^2}{x_2^2} + \frac{dx_3^2}{x_3^2} \; + \; 2\,\frac{x_1}{x_2}\,dx_4^2 \; + \; 2\,\frac{x_2}{x_3}\,dx_6^2 \; + \; 2\,\frac{x_1}{x_3}\,\big(dx_5 - x_6\,dx_4\big)^2.$$

Which, combined with the fact that in the spatial-feature space we use the Euclidean metric on the spatial part, we finally derive the final expression for the metric matrix

$$h = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{x_1^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{x_2^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{x_3^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2(\frac{x_1}{x_2} + \frac{x_6^2x_1}{x_3}) & \frac{-2x_6x_1}{x_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{-2x_6x_1}{x_3} & \frac{2x_1}{x_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2x_2}{x_3} \end{pmatrix}$$

From here we continue and derive the expression for the induced metric

$$\gamma_{\mu\nu}(x) = \partial_\mu X^i \partial_\mu X^j h_{ij}(X) = X_\mu^T h X_\nu$$

Which in general states

$$\gamma(x) = (D_xX)^T h(X) D_xX$$

6 Calculating the Euler-Lagrange equations of the model functional

The functional that we aim to minimize in this paper is given by

$$S(D,f) = \int_\Omega (||A_{bi\_tensor}(D,f) - \hat{A}||_2^2 + \alpha\sqrt{|\gamma(D)|})d\Omega$$

Where $|\gamma| = det(\gamma)$, α is the coefficient of the smoothness regularization term of the functional, Ω is the 3-dimensional spatial domain of the scan and $\hat{A}$ is the observed data produced by the diffusion MRI scanner. So, the Lagrangian of this functional is

$$L(X^1, X^2, \ldots, X^6, X_x^1, X_y^1, X_z^1, \ldots, X_x^6, X_y^6, X_z^6) = ||A_{bi\_tensor}(D,f) - \hat{A}||_2^2 + \alpha\sqrt{|\gamma(D)|}$$

Where $X^1, \ldots, X^6$ are the Iwasawa coordinates used in decomposing $\mathbf{D}$ into the form N^TAN seen in previous section.
And hence, for some $i \in [6]$ we have that the Euler-Lagrange equation is

$$\frac{\partial S}{\partial X^i} = \frac{\partial L}{\partial X^i} - \frac{\partial}{\partial x}(\frac{\partial L}{\partial X_x^i}) - \frac{\partial}{\partial y}(\frac{\partial L}{\partial X_y^i}) - \frac{\partial}{\partial z}(\frac{\partial L}{\partial X_z^i}) = 0$$

First, we look at the first part of the EL equations

$$\frac{\partial L}{\partial X^i} = \frac{\partial}{\partial X^i}(||A_{bi\_tensor}(D,f) - \hat{A}||^2 + \alpha\sqrt{|\gamma(D)|}) = \sum_{k=1}^{n} \frac{\partial}{\partial X^i}[A_{bi\_tensor}(D,f) - \hat{A}]_k^2 + \alpha\frac{\partial}{\partial X^i}\sqrt{|\gamma(D)|}$$

$$= -2bf\sum_{k=1}^{n}(A_{bi\_tensor}-\hat{A})_k(A_{tissue})_k(\frac{\partial}{\partial X^i}(q_k^T D q_k)) + \alpha\frac{1}{2}\frac{1}{\sqrt{|\gamma|}}|\gamma|(Trace(\gamma^{-1}\frac{\partial}{\partial X^i}\gamma)) =$$

$$-2bf\sum_{k=1}^{n}(A_{bi\_tensor}-\hat{A})_k(A_{tissue})_k(q_k^T\frac{\partial D}{\partial X^i}q_k) + \frac{\alpha}{2}\sqrt{|\gamma|}Trace(\gamma^{-1}(DX)^T\frac{\partial h}{\partial X^i}DX)$$

We then focus on the second part of the RHS and express it in Einstein notation

$$Trace(\gamma^{-1}(DX)^T\frac{\partial h}{\partial X^i}DX) = \gamma^{\mu\nu}\partial_\mu X^k\partial_\nu X^l\frac{\partial h_{kl}}{\partial X_i}$$

Now, we continue to the calculation of the second part of the EL equations. We notice first that

$$\frac{\partial L}{\partial X_\mu^i} = 0 + \alpha\frac{\partial}{\partial X_\mu^i}(\sqrt{|\gamma|}) = \frac{\alpha}{2}\sqrt{|\gamma|}Trace(\gamma^{-1}\frac{\partial\gamma}{\partial X_\mu^i})$$

Here we notice that

$$\frac{\partial\gamma}{\partial X_\mu^i} = \frac{\partial}{\partial X_\mu^i}((DX)^T hDX) = \frac{\partial(DX)^T}{\partial X_\mu^i}hDX + (DX)^T h\frac{\partial(DX)}{\partial X_\mu^i} =$$

$$2\frac{\partial(DX)^T}{\partial X_\mu^i}hDX = 2\delta_{\mu i}hDX$$

Where $\delta_{\mu i}$ is a matrix with all entries zero except the entry (μ, i) which is equal to 1. Putting all together and expressing in Einstein notation gives

$$\frac{\partial\sqrt{|\gamma|}}{\partial X_\mu^i} = \sqrt{|\gamma|}\gamma^{\nu\mu}h_{ij}\partial_\nu X^j$$

And hence

$$\frac{\partial}{\partial\mu}(\frac{\partial L}{\partial X_\mu^i}) = \frac{\partial}{\partial\mu}(\alpha\frac{\partial\sqrt{|\gamma|}}{\partial X_\mu^i}) = \alpha\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}h_{ij}\partial_\nu X^j)$$

And so finally we get

$$\frac{\partial S}{\partial X^i} = -2bf\sum_{k=1}^{n}(A_{bi\_tensor}-\hat{A})_k A_{tissue}(q_k^T\frac{\partial D}{\partial X^i}q_k) + \alpha(\frac{1}{2}\sqrt{|\gamma|}\gamma^{\mu\nu}\frac{\partial h_{kl}}{\partial X_i}\partial_\mu X^k\partial_\nu X^l - \partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}h_{ij}\partial_\nu X^j))$$

7 Calculating the Gradient Descent update step X_t^i

Now, we will denote

$$dS = (\frac{\partial S}{\partial X^1}, \ldots, \frac{\partial S}{\partial X^9})$$

And see that the Euler-Lagrange equations are just

$$dS = \vec{0}$$

The equations remain true when we multiply it from the left by $\frac{1}{\sqrt{|\gamma|}}h^{-1}$, namely we now have

$$\frac{1}{\sqrt{|\gamma|}}h^{-1}dS = \vec{0}$$

writing the above in Einstein notation once again yields

$$\frac{1}{\sqrt{|\gamma|}}h^{ij}dS_i = \frac{1}{\sqrt{|\gamma|}}h^{ij}\frac{\partial S}{\partial X^i}$$

We are now setting the update step of the Gradient Descent to be

$$X_t^i = -\frac{1}{\sqrt{|\gamma|}}(h^{-1}dS)_i = -\frac{1}{\sqrt{|\gamma|}}h^{ji}\frac{\partial S}{\partial X^j}$$

We will now calculate the above expression in detail.

$$-\frac{h^{ji}}{\sqrt{|\gamma|}}\frac{\partial S}{\partial X^j} = \frac{2bfh^{ij}}{\sqrt{|\gamma|}}\sum_{k=1}^{n}(A_{bi\_tensor} - \hat{A})_k A_{tissue}(\frac{\partial}{\partial X^j}(q_k^T D q_k))-$$

$$\alpha(\frac{1}{2}\gamma^{\mu\nu}\partial_\mu X^k \partial_\nu X^l h^{ij}\frac{\partial h_{kl}}{\partial X_j} - \frac{1}{\sqrt{|\gamma|}}h^{ij}\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}h_{j\ell}\partial_\nu X^\ell)$$

We now focus separately on the two expressions on the right. We begin by noticing that we can use the differentiation product rule and get that

$$h^{ij}\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}h_{j\ell}\partial_\nu X^\ell) = \partial_\mu(h^{ij}\sqrt{|\gamma|}\gamma^{\nu\mu}h_{j\ell}\partial_\nu X^\ell) - \frac{\partial h^{ij}}{\partial\mu}\sqrt{|\gamma|}\gamma^{\nu\mu}h_{j\ell}\partial_\nu X^\ell =$$

$$\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}h^{ij}h_{j\ell}\partial_\nu X^\ell) - \sqrt{|\gamma|}\gamma^{\nu\mu}\frac{\partial h^{ij}}{\partial\mu}h_{j\ell}\partial_\nu X^\ell$$

Now, notice that $h^{ij}h_{jl} = \delta_l^i$ and also a usefull formula for the derivative of the inverse of the metric h

$$0 = \partial_\mu I = \partial_\mu(h^{-1}h) = (\partial_\mu h^{-1})h + h^{-1}(\partial_\mu h)$$

And hence

$$\partial_\mu h^{-1} = -h^{-1}(\partial_\mu h)h^{-1}$$

So

$$\frac{\partial h^{ij}}{\partial\mu} = -h^{im}\frac{\partial h_{mp}}{\partial_\mu}h^{pj}$$

Here we should also notice that from chain rule we should get the full expression which is

$$\frac{\partial h^{ij}}{\partial\mu} = -h^{im}\frac{\partial h_{mp}}{\partial_\mu}h^{pj} = -h^{im}\frac{\partial h_{mp}}{\partial X^q}\partial_\mu X^q h^{pj}$$

and hence we get that

$$h^{ij}\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}h_{j\ell}\partial_\nu X^\ell) = \partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}\partial_\nu X^i) + \sqrt{|\gamma|}\gamma^{\nu\mu}h^{im}\frac{\partial h_{mp}}{\partial X^q}\partial_\mu X^q h^{pj}h_{j\ell}\partial_\nu X^\ell =$$

$$\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}\partial_\nu X^i) + \sqrt{|\gamma|}\gamma^{\nu\mu}h^{im}\frac{\partial h_{mp}}{\partial X^q}\partial_\mu X^q \partial_\nu X^p$$

Now, we are going to notice that we can rewrite the following quantity

$$\frac{\partial h_{mp}}{\partial X^q}\partial_\mu X^q\partial_\nu X^p = \frac{1}{2}(\frac{\partial h_{mq}}{\partial X^p}\partial_\mu X^p\partial_\nu X^q + \frac{\partial h_{mp}}{\partial X^q}\partial_\mu X^q\partial_\nu X^p) = \frac{1}{2}(\frac{\partial h_{mq}}{\partial X^p}\partial_\mu X^p\partial_\nu X^q + \frac{\partial h_{pm}}{\partial X^q}\partial_\mu X^q\partial_\nu X^p)$$

And since the indices μ, ν are also summed over and γ^{-1} is symmetric , we can thus re-write again and get that

$$\frac{\partial h_{mp}}{\partial X^q}\partial_\mu X^q\partial_\nu X^p = \frac{1}{2}(\frac{\partial h_{mq}}{\partial X^p}\partial_\mu X^q\partial_\nu X^p + \frac{\partial h_{pm}}{\partial X^q}\partial_\mu X^q\partial_\nu X^p) = \frac{1}{2}(\frac{\partial h_{mq}}{\partial X^p} + \frac{\partial h_{pm}}{\partial X^q})\partial_\mu X^q\partial_\nu X^p$$

Plugging in back and getting the equation

$$h^{ij}\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}h_{j\ell}\partial_\nu X^\ell) = \partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}\partial_\nu X^i) + \frac{1}{2}\sqrt{|\gamma|}\gamma^{\nu\mu}h^{im}(\frac{\partial h_{mq}}{\partial X^p} + \frac{\partial h_{pm}}{\partial X^q})\partial_\mu X^q\partial_\nu X^p$$

Now, look at the following term that appears in the equation for X_t^i

$$\frac{1}{2}\gamma^{\mu\nu}\partial_\mu X^k \partial_\nu X^l h^{ij}\frac{\partial h_{kl}}{\partial X_j}$$

And see that it can be equivalently rewritten as

$$\frac{1}{2}\gamma^{\mu\nu}h^{im}\frac{\partial h_{qp}}{\partial X_m}\partial_\mu X^q \partial_\nu X^p$$

Now, by putting everything back together, we get that

$$X_t^i = \frac{2bfh^{ij}}{\sqrt{|\gamma|}}\sum_{k=1}^{n}(A_{bi\_tensor}+\hat{A})_k A_{tissue}(\frac{\partial}{\partial X^j}(q_k^T D q_k)) -$$

$$\frac{\alpha}{\sqrt{|\gamma|}}\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}\partial_\nu X^i) + \frac{\alpha}{2}\gamma^{\nu\mu}h^{im}(\frac{\partial h_{mq}}{\partial X^p}+\frac{\partial h_{pm}}{\partial X^q}-\frac{\partial h_{qp}}{\partial X_m})\partial_\mu X^q \partial_\nu X^p =$$

$$\frac{2bfh^{ij}}{\sqrt{|\gamma|}}\sum_{k=1}^{n}(A_{bi\_tensor}-\hat{A})_k A_{tissue}(q_k^T\frac{\partial D}{\partial X^j}q_k) + \frac{\alpha}{\sqrt{|\gamma|}}\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}\partial_\nu X^i) + \alpha\Gamma^i_{qp}\gamma^{\nu\mu}\partial_\mu X^q \partial_\nu X^p$$

Where Γ^i_{qp} is a Christoffel symbol of the Levi-Chivita connection on the spatial-feature manifold with the metric h. The general formula for the Christoffel symbols of the Levi-Chivita connections is given by

$$\Gamma^k_{ij} = \frac{1}{2}\sum_l h^{kl}[\frac{\partial h_{jl}}{\partial X^i}+\frac{\partial h_{il}}{\partial X^j}-\frac{\partial h_{ij}}{\partial X^l}]$$

Or in Einstein notation

$$\Gamma^k_{ij} = \frac{1}{2}h^{kl}(\frac{\partial h_{jl}}{\partial X^i}+\frac{\partial h_{il}}{\partial X^j}-\frac{\partial h_{ij}}{\partial X^l})$$

NOTE:
The final Gradient Descent equations here differ from the ones that appear in the paper. One thing that should be noticed is that in the functional $S(D, f)$ the fidelity term and the smoothness regularization term live in different spaces, the first lives in a Euclidean space while the second, on our embedded manifold with the metric h. In a discussion about the topic it occurred to me that the fidelity term should be corrected in such a way that the Gradient Descent update of this term will not be dependent on the metric h (or it's inverse). Hence, the corrected Gradient Descent equations should be

$$X_t^i = \frac{2bf}{\sqrt{|\gamma|}}\sum_{k=1}^{n}(A_{bi\_tensor}-\hat{A})_k A_{tissue}(q_k^T\frac{\partial D}{\partial X^i}q_k) + \frac{\alpha}{\sqrt{|\gamma|}}\partial_\mu(\sqrt{|\gamma|}\gamma^{\nu\mu}\partial_\nu X^i) + \alpha\Gamma^i_{qp}\gamma^{\nu\mu}\partial_\mu X^q \partial_\nu X^p$$

8 Explicitly Calculating the Cristoffel symbols of the Levi-Civita connection

As seen in the previous section, the Cristoffel symbols of the Levi-Civita connection on the Spatial-Feature space are given by the formula

$$\Gamma^k_{ij} = \frac{1}{2}h^{kl}(\frac{\partial h_{jl}}{\partial X^i}+\frac{\partial h_{il}}{\partial X^j}-\frac{\partial h_{ij}}{\partial X^l})$$

We are now going to calculate them explicitly for $k = 4, 5, \ldots, 9$. As first step, we calculate the inverse metric matrix, h^{-1}.

We first remind that the explicit expression for the metric matrix h is given by

$$h = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{x_1^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{x_2^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{x_3^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2(\frac{x_1}{x_2} + \frac{x_6^2 x_1}{x_3}) & \frac{-2x_6 x_1}{x_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{-2x_6 x_1}{x_3} & \frac{2x_1}{x_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2x_2}{x_3} \end{pmatrix}$$

And hence by noticing that h is block diagonal, we take the inverse of each block and arrive at

$$h^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x_2^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & x_3^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{x_2}{2x_1} & \frac{x_2 x_6}{2x_1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{x_2 x_6}{2x_1} & \frac{x_3 + x_2 x_6^2}{2x_1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{x_3}{2x_2} \end{pmatrix}$$

Now we proceed to the calculation of the Crystoffel symbols

$$\Gamma^4_{ij} = \frac{1}{2} h^{4l} (\frac{\partial h_{jl}}{\partial X^i} + \frac{\partial h_{il}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^l}) = \frac{1}{2} h^{44} (\frac{\partial h_{j4}}{\partial X^i} + \frac{\partial h_{i4}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^4})$$

$$\Gamma^4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{x_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{x_1^2(x_3 + x_2 x_6^2)}{x_2 x_3} & \frac{x_1^2 x_6}{x_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{x_1^2 x_6}{x_3} & -\frac{x_1^2}{x_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Gamma^5_{ij} = \frac{1}{2} h^{55} (\frac{\partial h_{j5}}{\partial X^i} + \frac{\partial h_{i5}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^5}), \quad \Gamma^6_{ij} = \frac{1}{2} h^{66} (\frac{\partial h_{j6}}{\partial X^i} + \frac{\partial h_{i6}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^6})$$

$$\Gamma^5 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{x_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{x_2^2}{x_3} \end{pmatrix}, \quad \Gamma^6 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{x_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_6^2 & -x_6 x_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -x_6 x_1 & x_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_2 \end{pmatrix}$$

$$\Gamma^7_{ij} = \frac{1}{2} h^{7l} (\frac{\partial h_{jl}}{\partial X^i} + \frac{\partial h_{il}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^l}) = \frac{1}{2} h^{77} (\frac{\partial h_{j7}}{\partial X^i} + \frac{\partial h_{i7}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^7}) + \frac{1}{2} h^{78} (\frac{\partial h_{j8}}{\partial X^i} + \frac{\partial h_{i8}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^8})$$

$$\Gamma^7 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2x_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2x_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2x_1} & -\frac{1}{2x_2} & 0 & 0 & 0 & \frac{x_2x_6}{2x_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{x_2}{2x_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{x_2x_6}{2x_3} & -\frac{x_2}{2x_3} & 0 \end{pmatrix}$$

$$\Gamma^8_{ij} = \frac{1}{2}h^{8l}(\frac{\partial h_{jl}}{\partial X^i} + \frac{\partial h_{il}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^l}) = \frac{1}{2}h^{87}(\frac{\partial h_{j7}}{\partial X^i} + \frac{\partial h_{i7}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^7}) + \frac{1}{2}h^{88}(\frac{\partial h_{j8}}{\partial X^i} + \frac{\partial h_{i8}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^8})$$

$$\Gamma^8 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2x_1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{x_6}{2x_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{x_6}{2x_3} & -\frac{1}{2x_3} & 0 \\ 0 & 0 & 0 & 0 & -\frac{x_6}{2x_2} & \frac{x_6}{2x_3} & 0 & 0 & \frac{1}{2}(\frac{x_2x_6^2}{x_3} - 1) \\ 0 & 0 & 0 & \frac{1}{2x_1} & 0 & -\frac{1}{2x_3} & 0 & 0 & -\frac{x_2x_6}{2x_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(\frac{x_2x_6^2}{x_3} - 1) & -\frac{x_2x_6}{2x_3} & 0 \end{pmatrix}$$

$$\Gamma^9_{ij} = \frac{1}{2}h^{99}(\frac{\partial h_{j9}}{\partial X^i} + \frac{\partial h_{i9}}{\partial X^j} - \frac{\partial h_{ij}}{\partial X^9})$$

$$\Gamma^9 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2x_2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2x_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{x_6x_1}{x_2} & \frac{x_1}{2x_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{x_1}{2x_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2x_2} & -\frac{1}{2x_3} & 0 & 0 & 0 \end{pmatrix}$$

9 Calculating the Gradient Descent step for f_t

We remind here that the bi_tensor model is defined as

$$A_{bi\_tensor}(\mathbf{D}, f) = fA_{tissue}(\mathbf{D}) + (1-f)A_{water}$$

Hence

$$f_t = -\frac{\partial S}{\partial f} = -\frac{\partial L}{\partial f} = -\frac{\partial}{\partial f}(\sum_{k=1}^{n}[A_{bi\_tensor}(\mathbf{D}, f) - \hat{A}]_k^2) = -2\sum_{k=1}^{n}(A_{bi\_tensor} - \hat{A})_k \frac{\partial}{\partial f}(A_{bi\_tensor})_k =$$

$$-2\sum_{k=1}^{n}(A_{bi\_tensor} - \hat{A})_k (A_{tissue} - A_{water})_k$$