

# Multi-compartment model analysis in diffusion tensor imaging

Rashi Samur

2012

Rashi, S. (2012). Multi-compartment model analysis in diffusion tensor imaging. Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/50594>

<https://doi.org/10.32657/10356/50594>



# **MULTI-COMPARTMENT MODEL ANALYSIS IN DIFFUSION TENSOR IMAGING**

**RASHI SAMUR**

School of Computer Engineering

A thesis submitted to the Nanyang Technological University  
in partial fulfilment of the requirement for the degree of  
Master of Engineering

**2012**

## Acknowledgments

---

It gives me immense pleasure to acknowledge the tremendous contribution of my teacher and supervisor Prof. Vitali Zagorodnov towards the successful completion of my M.Eng. thesis. He not only taught me the basics of the subject of my thesis but also helped me out with the actual implementation of the computations involved in my work, whenever I experienced difficulty in solving the problem. Professor Vitali also encouraged me continuously during the entire period of my M.Eng. course, whenever I found the going difficult owing to work related pressures. No words would suffice to express my sense of gratitude towards him.

I would also like to express my sincere gratitude to my employers – Tata Consultancy Services and all my seniors in the company for giving me the opportunity to pursue my M.Eng. degree. It would not have been possible for me to complete this work without their support.

I would like to express my gratitude to the Department of Computer Engineering, Nanyang Technological University for providing me the facilities and the opportunity to pursue my M.Eng. degree. It would not be out of place here to express my sincere thanks to all my teachers, who have taught me the required courses and contributed to my development.

I would also like to acknowledge the great help offered to me by my classmates, who were always by my side when I needed their help.

No words can adequately express my sense of gratitude for Prof. Lalit Kumar Goel, from the School of Electronics & Electrical Engg., in counselling me, and providing support throughout my academic life at Nanyang Technological University.

And last but not the least; I would like to acknowledge the tremendous help and moral support provided by my parents during the entire duration of my M.Eng. study. This dissertation would not have materialized but for their contribution.

# Contents

---

Acknowledgments.....	1
Contents .....	2
List of Tables .....	4
List of Figures .....	5
Abstract .....	6
1 Introduction .....	7
1.1 Motivation .....	7
1.2 Objectives.....	9
1.3 Organisation of the Report .....	9
2 Literature Review .....	10
2.1 Significance of Motion in Brain Imaging .....	10
2.2 Region-wise Diffusion Behaviour.....	12
2.3 Measurement of Diffusion Coefficient .....	12
2.4 Diffusion Tensor Imaging .....	15
2.4.1 ADC (Average Diffusion Coefficient) Maps.....	16
2.4.2 Anisotropy Maps.....	18
2.4.3 Fibre Orientation Maps .....	20
2.5 Multiple Compartment Analyses Models in DWI .....	21
2.5.1 Intracellular/Extracellular Compartments.....	23
2.5.2 Vascular/Water Diffusion Compartments.....	24
3 Problem Statement and Solutions.....	26
3.1 Problem Statement .....	26
3.2 Proposed Solution .....	26
3.2.1 Data Pre-processing .....	27
3.2.2 Mono-exponential Model Fitting.....	27

3.2.3	Bi-exponential Model Fitting.....	29
3.2.4	Model Selection .....	32
3.2.5	Summary of the Solution .....	33
4	Implementation .....	34
4.1	Mono-exponential Model Fitting .....	34
4.2	Bi-exponential Model Fitting .....	34
4.3	Data Simulation for Testing .....	36
5	Cramer-Rao Lower Bound (CRLB) for $D_{vas}$ .....	38
5.1	Introduction to CRLB.....	38
5.2	Relevance of CRLB for the Estimation Methods Used .....	39
5.3	Derivation of CRLB for $D_{vas}$ .....	40
5.4	Observations from the Derived CRLB .....	46
6	Experiment Results and Discussion .....	47
6.1	Data Pre-processing Results .....	47
6.2	Model Selection Results for Simulated Data .....	48
6.3	Comparison with CRLB Values.....	50
6.4	Comparison with MATLAB “lsqnonlin” Function.....	51
6.5	Model Fitting Results on Real MRI Signal Data .....	52
7	Conclusions and Future Work .....	55
7.1	Conclusions .....	55
7.2	Future Work .....	56
	References.....	57

## List of Tables

---

Table 1. Difference between scalar maps and fibre orientation maps.....	16
Table 2. Typical ADC Values .....	17
Table 3. Different Values of Anisotropy Maps for Different Kinds of Diffusion .....	19
Table 4. ADC for Fast and Slow Compartments in Regions of Interest .....	23
Table 5. Typical Three Compartment Diffusion for Cats .....	24
Table 6. Partial Derivatives of the Residual Sum of Squares .....	30
Table 7. Estimation of $D_{vas}$ subject to Noise Level and Water Fraction .....	48
Table 8. Estimated Error vs. CRLB for Different Levels of $\sigma$ and $f_{wat}$ .....	50
Table 9. Comparison of Results with MATLAB Non-linear Fitting Function .....	51
Table 10. Root Mean Squared Errors of CSF vs. WM & GM .....	53
Table 11. Estimated parameters in Bi-exponential and Mono-exponential voxels.....	54
Table 12. The Ground Truth .....	54

## List of Figures

---

Figure 1. T1- and T2- Weighted Image[2] .....	7
Figure 2. Brain Map Generated by DTI .....	8
Figure 3. Motion of Patient.....	10
Figure 4. Difference between Flow and Diffusion .....	11
Figure 5. De-phase and Re-phase Experiment .....	13
Figure 6. Diffusion Effect on MR Signal .....	14
Figure 7. Flow Effect on MR signal .....	14
Figure 8. De-phasing and Re-phasing Time Length.....	15
Figure 9. Typical ADC image.....	17
Figure 10. Fractional Anisotropy Map .....	19
Figure 11. 24-bit DTI Colour Map .....	20
Figure 12. ADC Histograms at Different Maxima of b-Values .....	22
Figure 13. Brain with Skull .....	27
Figure 14. Bi-exponential Signal: $f(1 - bD_1) + (1 - f)e^{-bD_2}$ .....	36
Figure 15. Mono-exponential Signal: $e^{-bD_{mono}}$ .....	37
Figure 16. Distribution of Original Simulated Data .....	37
Figure 17. Skull Removal Result .....	47
Figure 18. Error Levels of $D_{vas}$ .....	49
Figure 19. Estimated Separation of Both Models with and without Noise.....	49
Figure 20. Original MRI Image vs. Number of Compartments Distribution .....	53

## **Abstract**

---

Diffusion Weighted Imaging (DWI) and Diffusion Tensor Imaging (DTI) are newly emerging techniques in Magnetic Resonance Imaging (MRI). These techniques enable studying connectivity and fibre orientations in different regions of the brain and also detecting abnormalities due to pathological conditions and physiological defects in the brain, which were not possible with conventional MRI techniques.

One challenging area of research in DWI is the estimation of multiple diffusion compartments within individual voxels of a DWI image. The concept behind this estimation is that different regions of the brain have different diffusivities, and a single voxel in the image can contain more than one such diffusion component.

This report presents our thorough study of multi-compartment estimation problem in DWI, focusing primarily on two compartment estimation problem due to two major types of diffusion compartments – water diffusion and vascular blood flow.

There are two methods of estimation proposed in this study, both achieving accurate results when compared with the ground truth values. The algorithms are tested both on synthetic data as well as on actual brain data.



# 1 Introduction

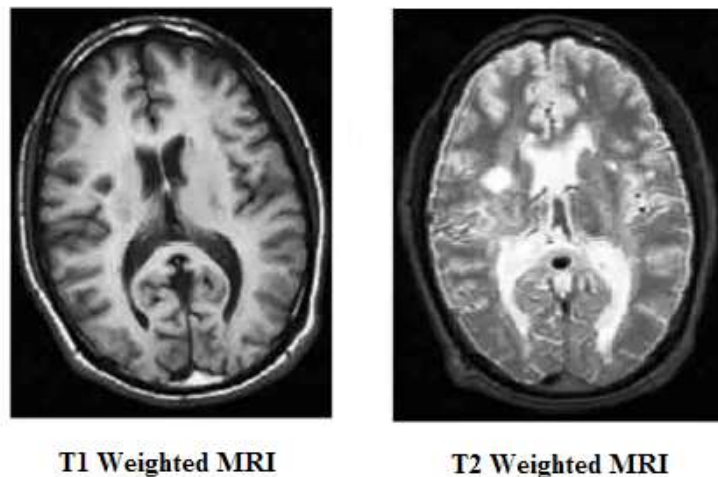
---

## 1.1 Motivation

The human brain can be divided into three main tissue regions: White Matter (WM), Gray Matter (GM), and Cerebrospinal Fluid (CSF). Gray Matter contains neurons which are locally inter-connected between each other. More distance connections between neurons occur via axonal fibres located within WM. Studying integrity of these connections is of interest because of their association with brain pathology.

The most common medical imaging technique, used in radiology, is Magnetic Resonance Imaging (MRI) [1]. This technique is used to visualise the internal structure and functions in the body, including the brain.

Conventional MRI methods like, T1- and T2- weighted images, show the 3 major regions of the brain using different light intensities. Examples of T1- and T2- weighted images are given in Figure 1.

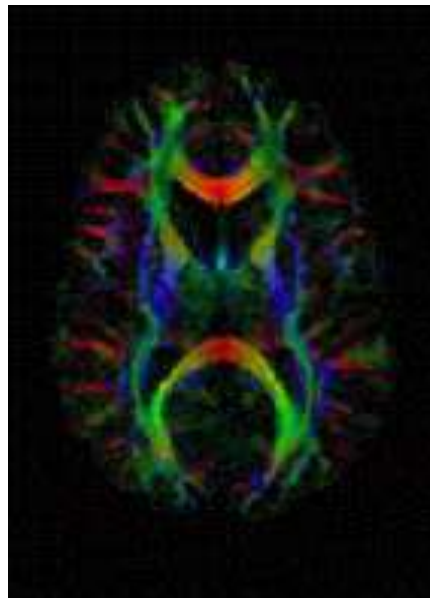


**Figure 1. T1- and T2- Weighted Image[2]**

In T1-weighted images, WM is displayed with the highest intensities, followed by the GM and lastly CSF with the darkest intensities. In T2-weighted images, it is the other way round, i.e. CSF region is the lightest and the WM region is the darkest.

The main disadvantage of conventional MRI is homogenous appearance of image intensities within a particular region and thus inability to reveal detailed anatomical structures within the different regions of the brain. This propels us to focus on the new emerging technique in the field of MRI – Diffusion Weighted Imaging (DWI). This technique takes the advantage of a very common phenomenon in biological tissues, i.e. diffusion of water molecules, and can bring to light the minute structural variations and anomalies within the tissue. It can be shown that water diffuses more rapidly in the direction aligned with the internal structure, such as the neural axons of the WM in the brain or muscle fibres in the heart, and more slowly in the perpendicular directions.

Hence, by measuring diffusivities in multiple directions the internal structure of the tissue can be well visualised. The DTI technique has enabled researchers to perform neural tractography (Figure 2 as example), to study the connectivity of different regions in the brain or to examine neural degeneration and demyelination, caused by traumas, tumours and inflammations.



**Figure 2. Brain Map Generated by DTI**

The study of DWI can be a breakthrough for many brain imaging problems. One of these problems to be discussed in this report is the estimation of diffusion properties in multiple brain compartments [3].

## **1.2 Objectives**

The main objectives of this research are twofold. The first is to understand the RWI/DTI method and its application to brain imaging. The second is to find a new model to estimate multiple compartments in DWI based brain images and at the same time estimate the volume fraction and diffusivity of each compartment in this new model.

## **1.3 Organisation of the Report**

This report is divided into six chapters. Chapter 1 introduces us to the DTI technique and the objectives of this research. Chapter 2 sheds more light on the DTI technique. Chapter 3 presents the imaging estimation problem in detail, and its proposed solutions. Chapter 4 discusses the MATLAB implementation of the algorithms presented in the previous chapter. Chapter 5 focuses on the CRLB estimation of the vascular diffusion coefficient  $D_{vas}$ . Chapter 6 contains experimental results and discussions on their significance, also with a comparison between the estimated values and the CRLB. Finally, Chapter 7 presents the conclusions of the proposed solution and the future work to be done.

## 2 Literature Review

---

### 2.1 Significance of Motion in Brain Imaging

Three types of motion have been identified in brain images, viz. bodily motion of patients during the scan, flow of body fluids and water molecular diffusion.

#### *Motion of Patient*

It is a very common scenario that during any kind of medical scans, the patients might move, which causes a bulk movement of water molecules in the body. As a result, the MRI images captured get corrupted, as shown in Figure 3. This kind of motion is not of interest to researchers and can actually hinder the visualisation of the other kinds of motion in the images.



**Figure 3. Motion of Patient**

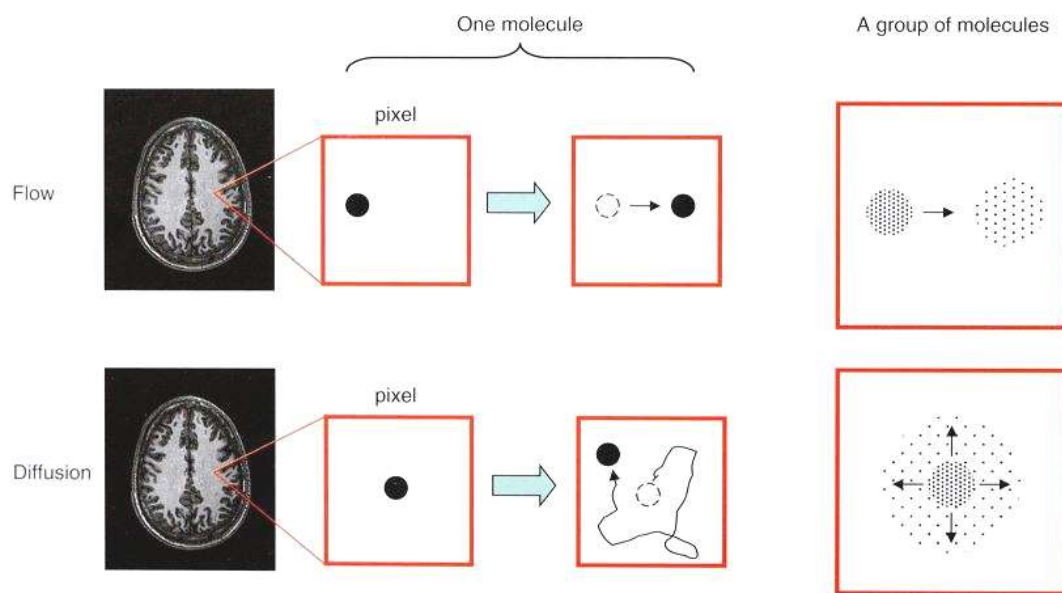
#### *Flow*

The second kind of motion – flow, is the unidirectional movement of a group of water molecules. This way the centre of this group changes position with the movement. Flow in the brain is typically caused by blood. The measurement of flow is complex and usually ignored, since water accounts for only 5% of volume in the blood vessels.

### *Diffusion*

This is the type of motion that draws our interest. Diffusion is a type of intra-voxel incoherent motion or random motion, also known as Brownian motion. Assuming there are no structural barriers to the molecular motion, the centre of the molecular group stays at the same place and the molecules spread out in random directions, following a Gaussian probability distribution.

The pixel size of MRI is normally 2mm x 2mm. The amount of flow is usually larger than the pixel size and hence very difficult to notice; but the actual amount of diffusion is 5-10  $\mu\text{m}$  approximately. So, in reference to the pixel size, the diffusion gets exaggerated.



**Figure 4. Difference between Flow and Diffusion**

When the diffusion is equal and unconstrained in all directions, it is called **isotropic**. But the micro-structure of brain tissues forms physical barriers that restrict the free movement of water molecules in some directions and promote in other directions. As a result, the diffusion becomes **anisotropic**.

## **2.2 Region-wise Diffusion Behaviour**

### ***Gray Matter (GM)***

The diffusion detected in DWI shows different characteristics in different regions. The image resolution for DWI is normally 2mm x 2mm, as clarified before. At this resolution, the fibre structures of the GM mostly appear incoherent, due to their structural complexity and relatively smaller size. The resulting water diffusion might also appear to be isotropic, although the speed of diffusion is actually constrained.

### ***White Matter (WM)***

In the WM, many axonal bundles are larger than the pixel size of the image, i.e. 2 mm. Hence, the diffusion will mostly appear anisotropic, following one dominant direction, according to the direction of the tissue structure.

### ***Cerebrospinal Fluid***

Lastly, the CSF region, being fluid, provides least constraint to the motion of water molecules and hence the diffusion is more or less isotropic.

## **2.3 Measurement of Diffusion Coefficient**

The next step is to define the Diffusion Coefficient (DC), as a parameter to measure diffusion along a particular direction. The higher the value of DC, the greater is the probability of the water molecules to change their location within a fixed amount of time. Hence, by measuring DC in different directions, the three main structures can be differentiated and studied.

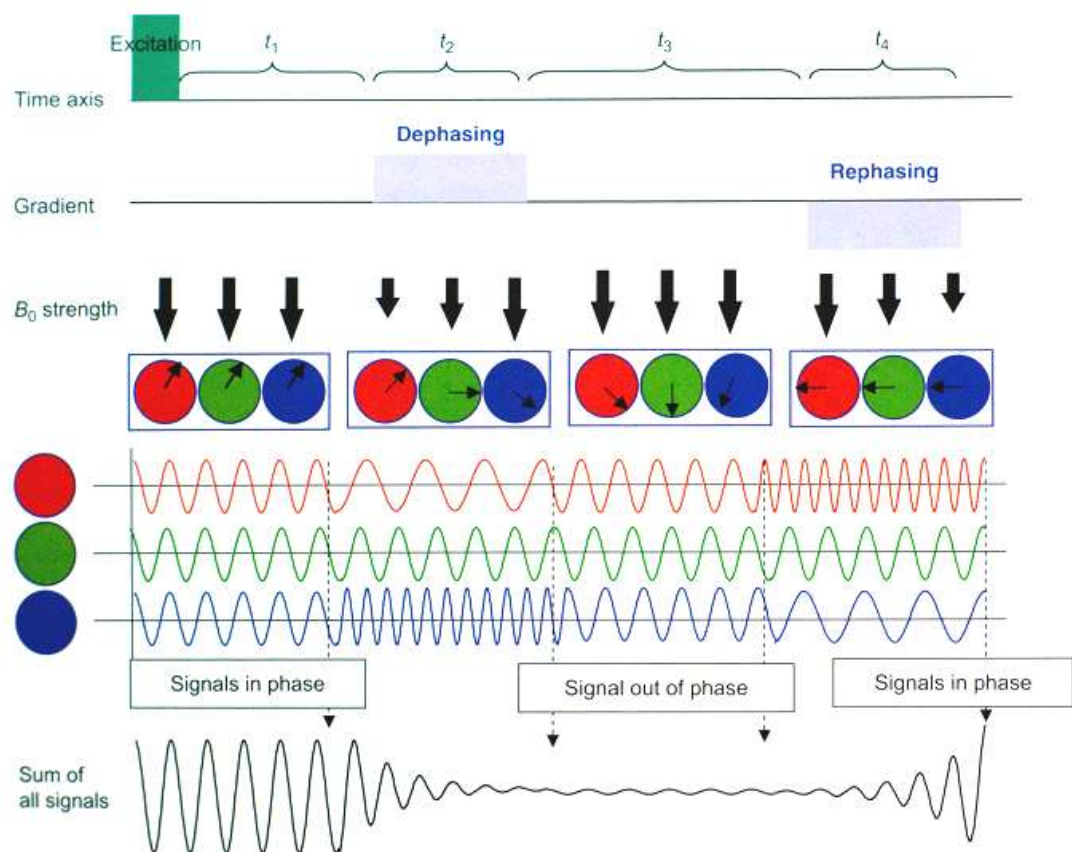
To measure DC in a direction, a de-phase/re-phase experiment is conducted. Assuming the water molecules in the brain are static, Figure 5 shows how the signals would behave.

### ***De-Phase***

When the de-phasing gradient is applied, different locations have varied magnetic strengths and hence the signal frequency for the different water molecules will also be different. Next, the same magnetic strength is applied to the whole and now the frequencies of all signals are the same. But, the signals are out of phase already.

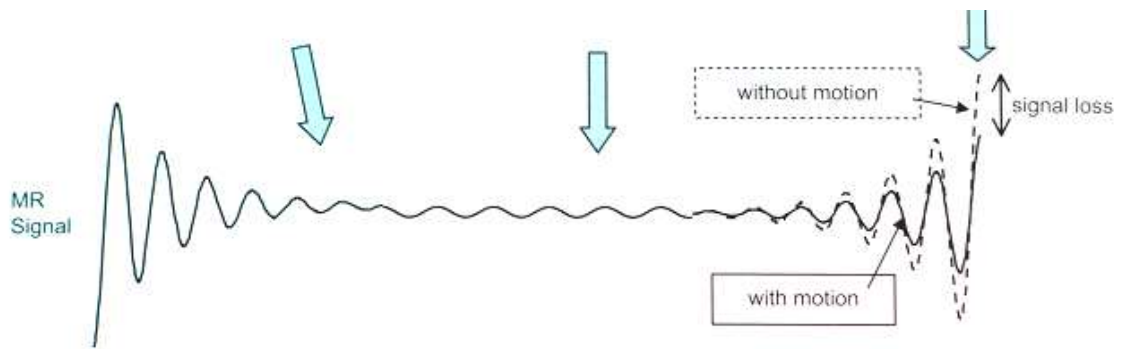
### ***Re-Phase***

After some time, the re-phasing gradient is applied to bring the signals back into phase. However, in this step, the water molecules might move around and cause phase disruption of the MR signals. This will result in signal loss, as shown in Figure 6.



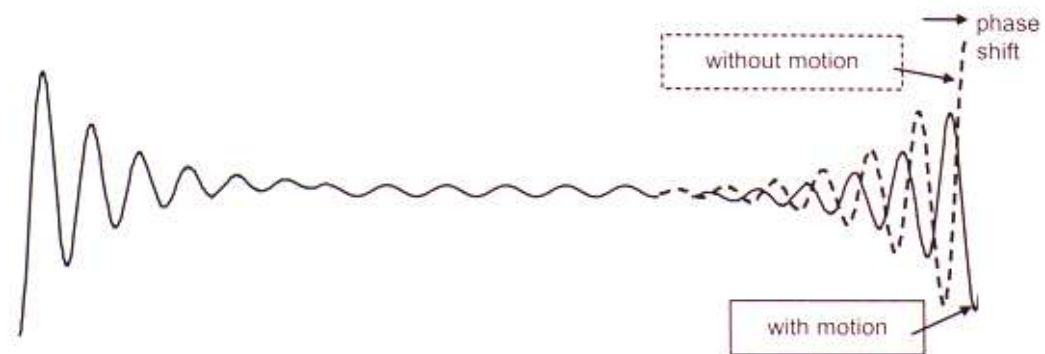
**Figure 5. De-phase and Re-phase Experiment**





**Figure 6. Diffusion Effect on MR Signal**

Another problem that surfaces in the process is that of phase shifting, due to flow motion happening between the de-phasing and re-phasing. This is shown in Figure 7.

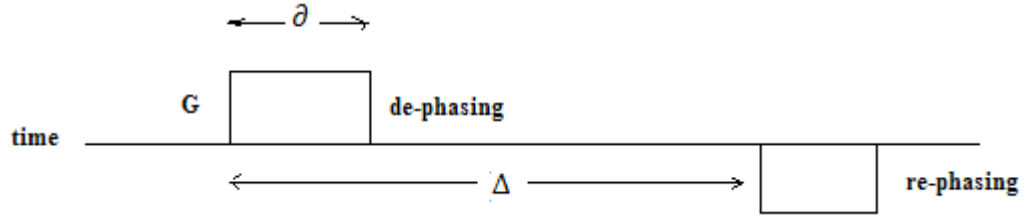


**Figure 7. Flow Effect on MR signal**

Since diffusion can induce signal loss, the measurement of diffusion controlling parameters becomes of high importance for this research. It appears that, the higher the DC or the longer the interval between de-phase and re-phase gradient, the more there is chance of water molecules moving around and causing higher signal loss. Also, the more the initial amount of de-phasing, determined by strength and time length, the more the diffusion causes phase disruption and hence signal loss.

Parameters defined as in Figure 8:  $G$  = de-phasing strength;  $\partial$  = de-phasing time length;  $D$  = diffusion coefficient; and  $\Delta$  = time difference between de-phasing and re-phasing gradients.





**Figure 8. De-phasing and Re-phasing Time Length**

Hence, the signal attenuation ratio  $= \frac{S}{S_0}$ , will be a function of  $G$ ,  $\partial$ ,  $\Delta$  and  $D$ . Since  $G$ ,  $\partial$ ,  $\Delta$  are known and the attenuation ratio can be measured through experiments, it is possible to calculate  $D$ , i.e. the diffusion coefficient.

Through mathematical characterization and experiments, an expression to calculate  $D$ , along the direction of the applied gradient, is derived as below [4]:

$$\frac{S}{S_0} = e^{-r^2 G^2 \partial^2 (\Delta - \partial/3) D} \text{ or } \frac{S}{S_0} = e^{-bD}, \text{ where } b = r^2 G^2 \partial^2 \left( \Delta - \frac{\partial}{3} \right) \text{ and } r \text{ is a constant.}$$

The above can be re-written as follows, to calculate  $D$ :

$$\ln S = \ln S_0 - bD \quad \text{or} \quad D = \frac{\ln S_0 - \ln S}{b}$$

$D$  can be more accurately estimated using the least squares fitting method, by using different gradient values. The gradients provide a set of  $b$  and  $S$  pair values.

The above method can only be used in case of isotropic diffusion, where the same DC applies along all directions and the resulting diffusion is spherical in shape. However, in case of anisotropic diffusion, DC has to be calculated along multiple dimensions. This brings us to the basis of Diffusion Tensor Imaging (DTI) that assumes that anisotropic diffusion follows the shape of an ellipsoid.

## 2.4 Diffusion Tensor Imaging

DTI is based on the assumption that anisotropic diffusion follows an ellipsoid (or rank-2 tensor) shape. To define this shape in a 3-D space, 6 parameters are required  $(\lambda_1, \lambda_2, \lambda_3, V_1, V_2, V_3)$ .  $V_1, V_2, V_3$  are the 3 orthogonal direction vectors that describe the spatial orientation, and  $\lambda_1, \lambda_2, \lambda_3$  are the respective lengths along each direction.

The same de-phase / re-phase experiments are conducted to compute these 6 parameters along the 6 axes, viz.  $x$  axis,  $y$  axis,  $z$  axis,  $xy$  axis ( $\sigma$  degree rotation from  $x$  to  $y$  axis),  $xz$  axis ( $\sigma$  degree rotation from  $x$  to  $z$  axis), and  $yz$  axis ( $\sigma$  degree rotation from  $y$  to  $z$  axis). The corresponding DC values are  $D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{yz}$  and  $D_{xz}$ . We can write DC in the matrix form as follows, where  $V_1, V_2, V_3$  are the Eigenvectors and  $\lambda_1, \lambda_2, \lambda_3$  are the Eigenvalues:

$$\begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}$$

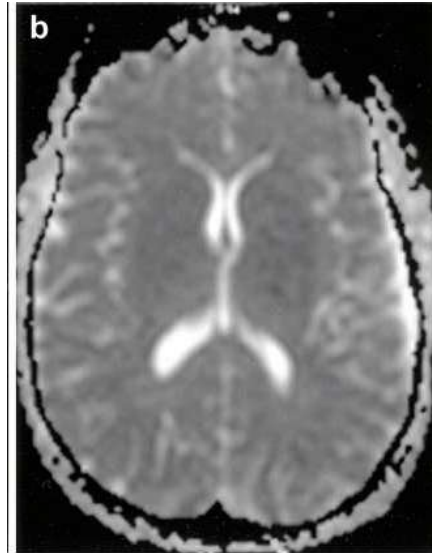
This DC matrix has to be determined at each pixel. But, it is impractical to visualize a 3-D ellipsoid at every pixel in an image, since the standard computer screen only provides 8-bit grayscale or 24-bit colour representations. Also, visualizing separate images of the 6 parameters at the same time is not an option. Hence, it is essential to represent the DC matrix as a single parameter at each pixel, which can be defined as a function of the parameters  $\lambda_1, \lambda_2, \lambda_3, V_1, V_2, V_3$ . Using the above methodology, 3 different types of DTI maps exist, and can be summarised as given in Table 1.

**Table 1. Difference between scalar maps and fibre orientation maps**

Map Type		Derived from
Scalar Maps (8-bit grayscale)	ADC (Average Diffusion Coefficient) Map	Eigen Values: $\lambda_1, \lambda_2, \lambda_3$
	Anisotropy Maps	
Fibre Orientation Map (24-bit colour)		Eigen Vectors: $V_1, V_2, V_3$

### 2.4.1 ADC (Average Diffusion Coefficient) Maps

Average Diffusion Coefficient, as the name suggests, represents the average diffusivity of each pixel, and is not sensitive to the orientation of the fibres. The mathematical form is  $\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$ . Figure 9 shows a sample ADC map.



**Figure 9. Typical ADC image**

An ADC map mainly shows the degree of constraint in diffusion. The higher the constraint, the lower is the corresponding ADC value. Since, fibres and myelin in the GM and WM regions apply greater constraint, the ADC values are much lower as compared to the CSF region, which is fluid and hence applies minimal constraint on the diffusion. Based on the ADC values, we can segregate CSF from WM and GM. But, between WM and GM, the values are still too close to be deterministic. Typical ADC values are listed here.

**Table 2. Typical ADC Values**

	ADC ( $\times 10^{-3} \text{ mm}^2/\text{s}$ )
CSF	$3.19 \pm 0.10$
Gray Matter (frontal lobe)	$0.88 \pm 0.04$
Gray Matter (putamen)	$0.73 \pm 0.03$
White Matter	$0.78 \pm 0.06$

The advantage of using ADC maps, over T1- and T2- weighted images, is that ADC maps can detect myelinisation, loss of myelin and axonal fibres, and decreased blood flow in the brain. These features prove very useful in research and clinical applications.

### 2.4.2 Anisotropy Maps

There are 3 sub-types of Anisotropy maps – Fractional Anisotropy (FA), Relative Anisotropy (RA) and Volume Ratio (VR). They are used to measure the level of isotropy at each voxel. [5]

#### ***Fractional Anisotropy (FA)***

This is the measure of the fraction of the diffusion tensor attributed to anisotropy. The FA value is directly proportional to the level of anisotropy.

$$FA = \sqrt{\frac{1}{2} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

#### ***Relative Anisotropy (RA)***

RA value is a ratio between the anisotropic and isotropic portions of the diffusion tensor. This value is also directly proportional to the level of anisotropy.

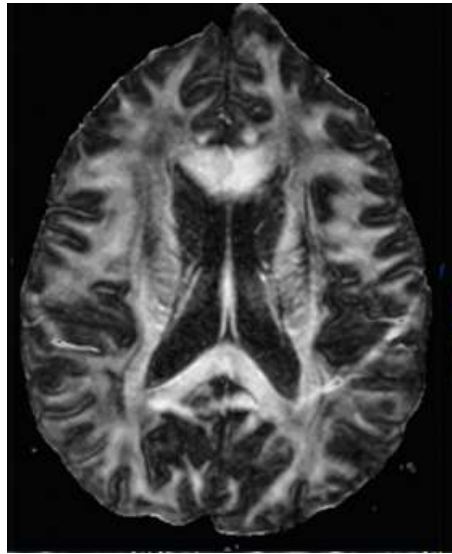
$$RA = \sqrt{\frac{1}{2} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\lambda_1 + \lambda_2 + \lambda_3}}$$

#### ***Volume ratio (VR)***

This value gives the relationship between the anisotropic diffusion ellipsoid and an ideal isotropic diffusion sphere. VR is inversely proportional to the level of anisotropy, making it different from the first two measures.

$$VR = \frac{\lambda_1 \lambda_2 \lambda_3}{\left(\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}\right)^3}$$

The most commonly used anisotropy map is FA. Figure 10 shows a sample.



**Figure 10. Fractional Anisotropy Map**

As explained before, due to the structural complexity of the GM fibres, the diffusion appears more isotropic and hence gives a lower FA value (dark pixels). In WM, due to myelinisation and fibres, the diffusion is more anisotropic and FA values are higher (bright pixels). For CSF, the diffusion is most isotropic and hence lowest values of FA, which result in the darkest pixels.

For anisotropic diffusion, the motion of water molecules can be dominant in one direction (tubular anisotropy), or across a plane in two directions (planar anisotropy). The different values of the 3 anisotropy maps for each of these diffusions can be summarised as below.

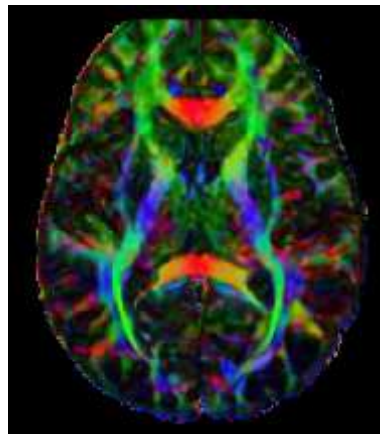
**Table 3. Different Values of Anisotropy Maps for Different Kinds of Diffusion**

	Isotropic ( $\lambda_1 = \lambda_2 = \lambda_3$ )	Tubular Anisotropy ( $\lambda_1 > \lambda_2 \approx \lambda_3$ )	Planar Anisotropy ( $\lambda_1 \approx \lambda_2 > \lambda_3$ )
FA	0	1	$\sqrt{1/2}$
RA	0	1	$1/2$
VR	1	0	0

### 2.4.3 Fibre Orientation Maps

These are also called Colour Maps. ADC and Anisotropy maps are scalar in nature, since they are derived from the scalar eigenvalues ( $\lambda_1, \lambda_2, \lambda_3$ ). Colour maps use the eigenvectors ( $V_1, V_2, V_3$ ) instead. These parameters represent orientation in 3-D space and hence are difficult to visualise in 2-D imaging. In DTI, the dominant vector  $V_1$ , which represents the local fibre orientation, is used, and the other two are discarded.

We normalise  $V_1$  to a unit vector,  $[x, y, z]$  (each scaled between 0 and 1) that satisfies the equality  $x^2 + y^2 + z^2 = 1$ . The resulting values are scaled up by 255 to give  $[255x, 255y, 255z]$  values, where 255x, 255y, 255z represent the Red, Green and Blue levels at each pixel, respectively.



**Figure 11. 24-bit DTI Colour Map**

E.g. If  $V_1$  is normalized to  $\left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right]$ , the RGB values would be  $[181, 181, 0]$ , and the resulting colour would be yellow.

It can be concluded that scalar maps complement conventional MRIs. To distinguish a region of interest, we can obtain values from T1 and T2 images, and ADC and FA maps. These values can be compared against population average values to determine the degree of abnormality. The advantage of DTI scalar maps over conventional MRI methods is that some structural abnormalities can be directly visualised in DTI images, e.g. diseases involving significant changes in the WM structure.

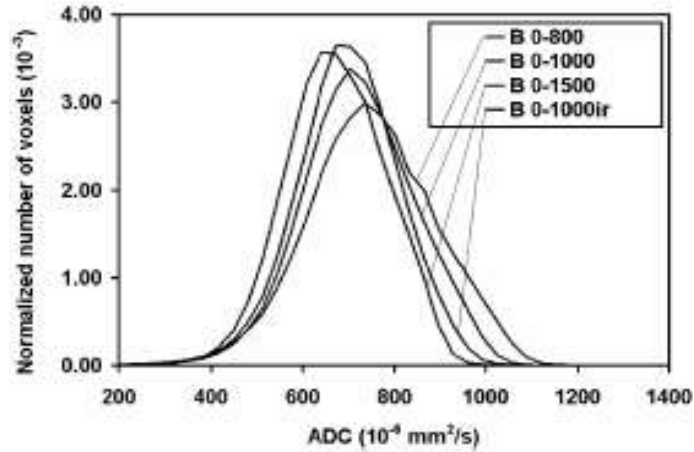
As for colour maps, the fibre orientations revealed cannot be achieved using conventional MRI. These maps can be used to study connectivity in the brain. They can also provide information about the size and shape of specific tracts in the nerve fibres.

Now we can see how important DTI has become in the area of neuroscience. As an emerging technique, DTI has a lot of new and hot research topics. And one of them is estimating multiple compartments in DTI. It is based on the assumption that in each voxel there is not only one diffusion compartment, but multiple compartments with different diffusion coefficients along one direction. Researchers try to find out these different compartments to get more information about the internal brain structures. Details of this topic and previous work are given in the next section.

## **2.5 Multiple Compartment Analyses Models in DWI**

As discussed in Section 2.3, the value of  $D$  (diffusion coefficient), along a single direction, can be determined using the signal attenuation value,  $\frac{S}{S_0} = e^{-bD}$ . In the ideal scenario, a single voxel, having a single homogenous compartment, will follow the above given exponential signal decay. This implies that, within a voxel, the  $DC$  value is the same everywhere, regardless of the direction. Hence,  $ADC$  values computed with different  $b$ -factors will remain unchanged (isotropic).

However, this model cannot be applied to the real-life scenario. Steens [6] conducted an experiment to study the effect of different maxima of  $b$ -values on  $ADC$  histograms. Multi-slice, single shot DW-EPI was performed in three dimensions for 7 healthy subjects, with  $b$ -values 0-800, 0-1000 and 0-1500. The resulting  $ADC$  histograms are shown in Figure 12.



**Figure 12. ADC Histograms at Different Maxima of b-values**

From the above figure, it is observed that the modes of the histograms shift towards the left and the mean ADC values decrease, as maximum b-values increase.

The concept of multi-compartments explains the decreasing ADC values over increasing b-values. There can be several compartments, with different ADCs, occupying a single voxel, sharing volume fractions, between which no exchange of water molecules happens. The mathematical form is illustrated here as:

$$\text{Signal Decay} = \sum_{i=1}^n f_i e^{-bD_i}$$

E.g., if a two-compartment model, with diffusion coefficients  $D_1$  and  $D_2$ , is taken as a single compartment, the resulting diffusion coefficient will follow  $e^{-bD} = f_1 e^{-bD_1} + f_2 e^{-bD_2}$ , and  $D$  will not be a constant value, varying according to b-value.

Based on the above concept, two models are defined. Details are given in further sections:

**Intracellular / Extracellular**, considering that the diffusivity inside and outside the cells are very different. The resultant ADC takes into account both compartments.

**Vascular / Water diffusion**, which considers the blood flow, along with the motion of water molecules in the tissue. The resultant MR signal is a sum of both vascular motion and water diffusion.



### 2.5.1 Intracellular/Extracellular Compartments

In DWI, over extended range of b-factors, the signal decay deviates significantly from the basic mono-exponential model. One of the major reasons for this behaviour is attributed to the variance in diffusivities, between extracellular (faster diffusion) and intracellular (slower diffusion).

If the b-value is small, the time for diffusion to occur is very short and is not enough for the slower intracellular compartments to show any detectable diffusion. In this case, the diffusion motion is dominated by the extracellular compartments, which are faster. When the b-value is very high, the intracellular compartments display higher diffusion and become dominant over the extracellular space.

To investigate this behaviour, a very wide range of b-values (0 to 5000s/mm<sup>2</sup>) is used, with a simplified mathematical model [7]:

$$\text{Signal Decay} = f_f e^{-D_f * b} + f_s e^{-D_s * b}, \quad \text{where } (f_s + f_f = 1) \text{ and}$$

$D_f, D_s$  are ADC values, and  $f_f, f_s$  are volume fractions for fast and slow components, respectively.

The table below is a list of typical ADC values and volume fractions for fast and slow compartments.

**Table 4. ADC for Fast and Slow Compartments in Regions of Interest**

<b>Tissue</b>	<b><math>D_f</math> (<math>\mu\text{mm}^2/\text{ms}</math>)</b>	<b><math>D_s</math> (<math>\mu\text{mm}^2/\text{ms}</math>)</b>	<b><math>f_s</math></b>
Cortical GM next to CSF	1.479 $\pm$ 0.166	0.466 $\pm$ 0.017	0.510 $\pm$ 0.012
Cortical GM next to WM	1.142 $\pm$ 0.106	0.338 $\pm$ 0.027	0.378 $\pm$ 0.038
Thalamus	1.320 $\pm$ 0.164	0.271 $\pm$ 0.040	0.383 $\pm$ 0.069
Frontal WM	1.155 $\pm$ 0.046	0.125 $\pm$ 0.014	0.301 $\pm$ 0.050
Internal capsule WM	1.215 $\pm$ 0.024	0.183 $\pm$ 0.009	0.363 $\pm$ 0.020

## 2.5.2 Vascular/Water Diffusion Compartments

As discussed earlier, MR signal loss is a combination of both diffusion of water molecules and blood flow in the tissue. At lower b-values, between 200-300s/mm<sup>2</sup>, vascular blood flow and extracellular water molecules dominate the signal loss, since intracellular diffusion is negligible. Especially during neuronal activity, when blood flow becomes very fast, with small b-values, the ADC values increase a lot.

Tao Jin et al [7] proposed a 3-compartment model to investigate the vascular components in the animal brain, the 3 compartments being – tissue water, arterial blood pool and venous blood pool. For simplicity, the signal attenuations of both arterial and venous blood pools are assumed to be mono-exponential. The DC values of each compartment are listed in the table below, based on measurements from 13 female adolescent cats.

**Table 5. Typical Three Compartment Diffusion for Cats**

Parameter	Description	Value (μmm <sup>2</sup> /s)
D	diffusion coefficient of tissue water	0.8
D <sub>a</sub>	Pseudo-diffusion coefficient of arterial blood	20-150 (default 100)
D <sub>b</sub>	Pseudo-diffusion coefficient of venous blood	10-40 (default 20)

In 2009, E. Sala et al [8] have suggested a simpler estimation of vascular signals, based on an assumption that vascular flow can create significant signal loss only for b = 0 s/mm<sup>2</sup> image. Any images with b ≥ 100 s/mm<sup>2</sup>, will have negligible effects from vascular components, and can be purely attributed to water diffusion.

A mono-exponential model can be used to fit the portion of b ≥ 100 s/mm<sup>2</sup> of the signals, to extrapolate the intensity at b = 0 s/mm<sup>2</sup>, defined as S<sub>extrap</sub>. This value is assumed to be the water diffusion component at b = 0 s/mm<sup>2</sup>. The vascular fraction  $f_{vas}$  can then be computed as:

$$f_{vas} = \frac{S_0 - S_{extrap}}{S_0} = 1 - \frac{S_{extrap}}{S_0}$$

Based on the above defined multi-compartment models for DTI, the objective of this research is to estimate the diffusion coefficients and volume fractions of the water diffusion and vascular blood flow components. Since, blood flow in the brain is directly related to the brain activity (regions of interest), it is essential to include the understanding of vascular blood flow in this study.

## 3 Problem Statement and Solutions

---

### 3.1 Problem Statement

The ultimate goal of our approach is to be able to correctly estimate the vascular compartments and water diffusion compartments in DTI. To simplify the model, we do not distinguish between arterial and venous blood structures, or between GM, WM and CSF. Hence, effectively there are only two components – vascular flow with arterial and venous blood, and water diffusion in GM, WM and CSF.

The mathematical model for signal decay in multiple compartments was defined as:

$$\text{Signal Decay} = \sum_{i=1}^n f_i e^{-bD_i}$$

Hence, for the two-compartment model defined above:

$$\text{Signal Decay} = f_{wat} e^{-bD_{wat}} + f_{vas} e^{-bD_{vas}},$$

where  $f_{wat}$  = volume fraction of water component,  $f_{vas}$  = volume fraction of vascular component, and  $f_{wat} + f_{vas} = 1$ .

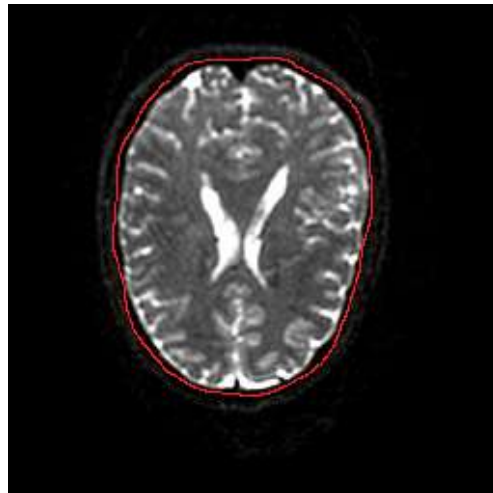
The next section describes the solutions proposed for estimating the above parameters of volume fractions and the exponentials involved to define the model.

### 3.2 Proposed Solution

To get more accurate estimations, data is pre-processed before applying the estimation methods. Voxel-by-voxel analysis is used to fit the measured signal into predefined mathematical models. The signal attenuation, in each voxel of the brain image, can be assumed to be a weighted sum of the vascular and the water diffusion components. The problem at hand is to segregate these two signals at each voxel.

### **3.2.1 Data Pre-processing**

The first step in pre-processing data is to remove the skull and scalp from the MR image. The skull is a bony structure that covers the brain of most animals. It supports the structure of the face and protects the brain against injury. Due to the structure of the skull, MR signals are produced and appear in the ADC maps. The ADC values are much smaller than the regions of GM, WM and CSF. In the presence of the skull structure in the image, it is difficult to apply correct estimation of the slow and fast diffusion components. The sample image below shows the skull located outside the red line.



**Figure 13. Brain with Skull**

For removal of the skull, a threshold value is applied at each pixel, equal to the minimum ADC value measured at different b-values. Anything below the threshold is considered as skull or background and is removed from the image.

### **3.2.2 Mono-exponential Model Fitting**

In order to compare the performances of the two-compartment model with the traditional mono-exponential model, several methods are developed. These methods fit MR signal data into the mono-exponential model.

The signal decay values  $y_i$  and b-values  $b_i$  are fitted into the mono-exponential curve:

$$y_i = e^{-b_i D_{mono}}$$

The above curve can be re-written in linear form as follows:

$$\log(y_i) = -b_i D_{mono}$$

For curve-fitting, the Residual Sum of Squares (RSS) is calculated as:

$$RSS = \sum_{i=1}^n [\log(y_i) + b_i D_{mono}]^2$$

The RSS needs to be minimised as below to get the optimum value of  $D_{mono}$ :

$$D_{mono} = \frac{\sum -\log(y_i) b_i}{\sum b_i^2}$$

Another approximation of the exponential curve is a straight line given by the below equation:

$$y_i = \exp(-b_i D_{mono}) \approx 1 - b_i D_{mono} \text{ (when } b_i \text{ are small)}$$

In this case, the RSS is as given below:

$$RSS = \sum_{i=1}^n (y_i - 1 + b_i D_{mono})^2$$

Minimising RSS gives the optimum value of  $D_{mono}$  as below:

$$D_{mono} = \frac{\sum (1 - y_i) b_i}{\sum b_i^2}$$

The above two approaches are quite similar to each other and hence give consistent results.

### 3.2.3 Bi-exponential Model Fitting

This method is applied at the individual voxel level. A simplified two-compartment model is used for each voxel, to evaluate the vascular and water diffusion components.

The two-compartment model equation can be written as:

$$\text{Signal Decay} = f_{\text{wat}}e^{-bD_{\text{wat}}} + f_{\text{vas}}e^{-bD_{\text{vas}}},$$

It is known that,  $0 < D_{\text{wat}}, D_{\text{vas}} \ll 1$  and  $D_{\text{wat}} < D_{\text{vas}}$ . Also,  $f_{\text{wat}} + f_{\text{vas}} = 1$ . Hence, the above can be further simplified as:

$$\text{Signal Decay} = f_{\text{wat}}(1 - bD_{\text{wat}}) + (1 - f_{\text{wat}})e^{-bD_{\text{vas}}}$$

The Residual Sum of Squares (RSS) in this model is also minimized to estimate the three unknown parameters ( $D_{\text{wat}}, f_{\text{wat}}, D_{\text{vas}}$ ) for each voxel. To reduce the above expression to a single estimable parameter, the following lemma is used. The proof of the lemma is given below:

#### **Lemma:**

$D_{\text{wat}}$  and  $f_{\text{wat}}$  can be expressed in terms of  $D_{\text{vas}}$ , for obtaining the minimum RSS.

#### **Proof:**

Recall our simplified two-compartment model:

$$y_i = f(1 - b_iD_1) + (1 - f) \exp(-b_iD_2), \quad i = \overline{1:n},$$

where  $n$  is the total number of b-values,  $f(f_{\text{wat}})$  is the water volume fraction,  $D_1(D_{\text{wat}})$  is the water diffusion component and  $D_2(D_{\text{vas}})$  is the vascular component.

The RSS can be written as:

$$RSS = \sum_{i=1}^n [y_i - f(1 - b_i D_1) - (1 - f) \exp(-b_i D_2)]^2$$

**Table 6. Partial Derivatives of the Residual Sum of Squares**

$\frac{\partial(RSS)}{\partial D_1} = 2 \sum_{i=1}^n [y_i - f(1 - b_i D_1) - (1 - f) \exp(-b_i D_2)] f b_i = 0$
$\frac{\partial(RSS)}{\partial D_2} = 2 \sum_{i=1}^n [y_i - f(1 - b_i D_1) - (1 - f) \exp(-b_i D_2)] (1 - f) b_i \exp(-b_i D_2) = 0$
$\frac{\partial(RSS)}{\partial f} = 2 \sum_{i=1}^n [y_i - f(1 - b_i D_1) - (1 - f) \exp(-b_i D_2)] [b_i D_1 + \exp(-b_i D_2) - 1] = 0$

We solve the following system of equations for the critical points of the function  $RSS(f, D_1, D_2)$  to find the minimum.

$$\begin{cases} \frac{\partial(RSS)}{\partial D_1} = 0 \\ \frac{\partial(RSS)}{\partial D_2} = 0 \\ \frac{\partial(RSS)}{\partial f} = 0 \end{cases}$$

Considering the partial derivative of  $D_1$ :

$$\frac{\partial(RSS)}{\partial D_1} = 0 \Rightarrow$$

$$\frac{1}{f} = \frac{\sum [\exp(-b_i D_2) - 1 + D_1 b_i] b_i}{\sum [\exp(-b_i D_2) - y_i] b_i} = a + b D_1,$$

$$\text{where } a = \frac{\sum [\exp(-b_i D_2) - 1] b_i}{\sum [\exp(-b_i D_2) - y_i] b_i}; b = \frac{\sum b_i^2}{\sum [\exp(-b_i D_2) - y_i] b_i}$$

Considering the partial derivative of  $D_2$ :



$$\frac{\partial(RSS)}{\partial D_2} = 0 \Rightarrow$$

$$\frac{1}{f} = \frac{\sum[\exp(-b_i D_2) - 1 + D_1 b_i] b_i \exp(-b_i D_2)}{\sum[\exp(-b_i D_2) - y_i] b_i \exp(-b_i D_2)} = c + d D_1,$$

$$\text{where } c = \frac{\sum[\exp(-b_i D_2) - 1] b_i \exp(-b_i D_2)}{\sum[\exp(-b_i D_2) - y_i] b_i \exp(-b_i D_2)}; d = \frac{\sum b_i^2 \exp(-b_i D_2)}{\sum[\exp(-b_i D_2) - y_i] b_i \exp(-b_i D_2)}$$

Combining the above two equations and solving for  $D_1$  gives:

$$D_1 = \frac{a-c}{d-b}, \quad \text{where } a, b, c, d \text{ are all functions of } D_2.$$

To calculate the value of  $f$ :

$$\frac{\partial(RSS)}{\partial f} = 0 \Rightarrow$$

$$f = \frac{\sum[\exp(-b_i D_2) - 1 + D_1 b_i][\exp(-b_i D_2) - y_i]}{\sum[\exp(-b_i D_2) - 1 + D_1 b_i]^2}$$

Hence, it is proved that  $f$  can be derived in terms of  $D_1$  and  $D_2$ , where  $D_1$  in turn can be expressed completely using  $D_2$ ; therefore  $f$  can also be derived from  $D_2$ . This reduces the 3-parameter estimation problem to the one that involves estimating just one parameter –  $D_2$ .

Another approach for bi-exponential curve fitting is to use the second partial derivatives of each of these parameters. But, due to their high noise sensitivities, these derivatives are impractical for experimentation purposes.

Thus, the solution for the bi-exponential model fitting is the Exhaustive Search method to find the value of  $D_2$  ( $D_{vas}$ ) for each voxel and to calculate the minimum square error or Residual Sum of Squares (RSS) using the derived values of the 3 parameters.

### 3.2.4 Model Selection

Different models suit different regions of the brain better. In some regions, the mono-exponential model fits better, e.g. the CSF region. In such cases, applying the bi-exponential model would yield inaccurate results. Therefore, the number of compartments should be decided for each voxel before finalizing the results of each parameter.

The error of fitting the mono-exponential model ( $RSS_1$ ) is greater than that of the bi-exponential model ( $RSS_2$ ), since the bi-exponential model is more flexible, due to more number of defining parameters. However, the cost of adding two more parameters,  $f$  and  $D_2$ , decreases the number of degrees of freedom in the model. Hence, it cannot be generalised that the second model will show more improvement than the first or vice-versa.

To compare the effect of the two factors – losing degrees of freedom ( $df$ ) vs. improving the RSS, on the level of performance, to decide a trade-off between the two models, a standard F-test statistical technique is applied [9]. The F-test is performed as below:

$$F = \frac{(RSS_1 - RSS_2)/(df_1 - df_2)}{RSS_2/df_2},$$

where the subscript 1 refers to the mono-exponential model and the subscript 2 refers to the bi-exponential model.

F-values are compared with a value  $F_0$  from the standard F-table [10], to decide which model fits better. If  $F \leq F_0$ , then model 1 is chosen, else model 2 is selected.

$$F \leq F_0 \leftrightarrow A \times RSS_1 - RSS_2 \leq 0,$$

$$\text{where } 0 \leq A = \frac{1}{1 + \frac{F_0(df_1 - df_2)}{df_2}} \leq 1$$

The above is equivalent to applying a threshold of zero on the weighted difference:  $(A \times RSS_1 - RSS_2)$ , to decide on the better model. If the difference  $\leq 0$ , mono-exponential model is selected, else bi-exponential model is selected.

In case of this particular problem, the parameters are as follows:

$$\begin{aligned} N = df &= 151 \text{ (151 distinct } b \text{ - values)} \\ \Rightarrow df_1 &= 151 - 1 = 150; df_2 = 151 - 3 = 148 \\ F_0 &= 3.0572; A \approx 0.9603 \end{aligned}$$

$F_0$  is looked-up from upper critical values of F-Distribution for 2-numerator degree of freedom and 148-denominator degree of freedom with 5% significance level.

### 3.2.5 Summary of the Solution

Following is a summary of the Voxel-by-Voxel fitting procedure described above:

1. Skull voxels are removed using thresholding techniques and categorised as background.
2. For each of the foreground voxels, attenuation signal fitting is done using both mono- and bi-exponential models.
3. The correct model is determined using the F-test on the effect of the two factors – degrees of freedom and residual sum of squares.
4. If F-test suggests mono-exponential (one-compartment) model, the equation is:

$$\text{Signal Decay} = e^{-bD}$$

5. Else, the bi-exponential (two-compartment) model equation is used as below:

$$\text{Signal Decay} = f_{wat}(1 - bD_{wat}) + (1 - f_{wat})e^{-bD_{vas}}$$

## 4 Implementation

---

The following section gives a brief explanation of how the key steps of the solution, discussed in Chapter 3, are implemented using MATLAB programming. Also, this chapter presents the data simulation and testing.

### 4.1 Mono-exponential Model Fitting

Referring to the two approaches for mono-exponential model fitting in Section 3.2.2, the MATLAB function implementations are as follows:

```
function [d_mono, lse] = mono_1(y, b)
d_mono = sum(b(:) .* (1 - y(:))) / sum(b(:).^2) * 10^3;
lse = sum((y(:) - 1 + d_mono * b(:) * 10^-3).^2);
```

```
function [d_mono, lse] = mono_2(y, b)
d_mono = sum(b(:) .* (-log(y(:)))) / sum(b(:).^2) * 10^3;
lse = sum((y(:) - exp(-d_mono * b(:) * 10^-3)).^2);
```

### 4.2 Bi-exponential Model Fitting

Referring to formulae derived for the three parameters –  $f$  and  $D_1$  from  $D_2$ , in Section 3.2.3, the following MATLAB function is defined to estimate the value of  $D_1$ , given the value of  $D_2$  and the input signals.

```
function d1 = d1_by_d2(d2, y, bValue)
ex = exp(-d2 * 10^-3 * bValue(:));
temp1 = sum(bValue(:) .* (ex - y(:)));
temp2 = sum((bValue(:) .* (ex)) .* (ex - y(:)));
a = sum((ex - 1) .* bValue(:));
a = a / temp1;
b = sum(bValue(:).^2);
b = b / temp1;
c = sum((bValue(:) .* (ex)) .* (ex - 1));
c = c / temp2;
d = sum((bValue(:).^2) .* ex);
d = d / temp2;
d1 = (a - c) / (d - b) * 10^3;
```

Similarly, for the derivation of parameter  $f$  from  $D_1$  and  $D_2$ , the MATLAB function defined is:

```
function f = f_by_d1_d2(d1, d2, y, bValue)
ex = exp(-d2 * 10^-3 * bValue(:));
ex2 = ex - 1 + d1 * 10^-3 * bValue(:);
temp1 = sum((ex2).^2);
f = sum((ex2) .* (ex - y(:)));
f = f / temp1;
```

To derive the value of  $D_2$ , an Exhaustive Search algorithm is deployed that gives the least square error fitting with the data at each voxel. The MATLAB function defined is given below.

```
function [a_f, a_d1, a_d2, LSE] = ES(y, bValue)
d2_end = 2000;
LSE = 10000000;
a_f = 0;
a_d1 = 0;
a_d2 = 0;

for d2 = 0 : 1 : d2_end
    d1 = d1_by_d2(d2, y, bValue); %Calculate d1 from d2 value
    f = f_by_d1_d2(d1, d2, y, b); %Calculate f from d2 and newly obtained d1
    SE = sum((y(:) - f * exp(-d1 * bValue(:) * 10^-3) - (1 - f) * exp(-d2 *
    bValue(:) * 10^-3)).^2);

    if (SE <= LSE) %Update minimal of Square Error function
        a_f = f;
        a_d1 = d1;
        a_d2 = d2;
        LSE = SE;
    end
end

%Ensure d1 < d2, if not swap d1 and d2 and change f to 1 - f
if(a_d1 > a_d2)
    temp = a_d1;
    a_d1 = a_d2;
    a_d2 = temp;
    a_f = 1 - a_f;
end
```

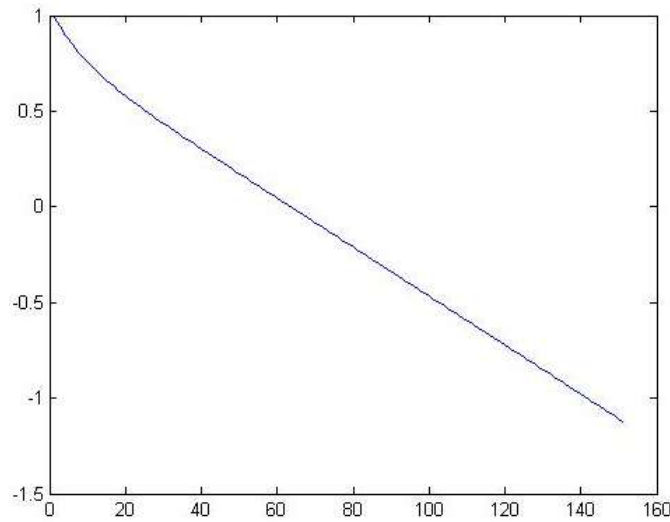
### 4.3 Data Simulation for Testing

The water diffusion component  $D_1$  is assumed to be  $1 \times 10^{-3} \text{mm}^2/\text{s}$ , and the vascular diffusion component  $D_2$  is assumed to be  $7 \times 10^{-3} \text{mm}^2/\text{s}$ . Based on these assumptions two sets of simulated data, one following the bi-exponential model, and the other following the mono-exponential model, were generated randomly.

In the first data set, synthetic water and vascular component coefficients are drawn from two normal Gaussian distributions, respectively, with  $\mathbf{D}_{1(10^{-3} \text{mm}^2/\text{s})} \sim \mathcal{N}(1, 1)$  and  $\mathbf{D}_{2(10^{-3} \text{mm}^2/\text{s})} \sim \mathcal{N}(7, 1)$ . Water volume fraction  $\mathbf{f}$  is iterated from 0.80 to 0.90. The pure signal decays are then computed from these synthetic parameters using the bi-exponential model equation:

$$\text{Signal Decay} = f(1 - bD_1) + (1 - f)e^{-bD_2},$$

where N distinct b-values = [0:16:2400] s/mm<sup>2</sup> are used.



**Figure 14. Bi-exponential Signal:  $f(1 - bD_1) + (1 - f)e^{-bD_2}$**

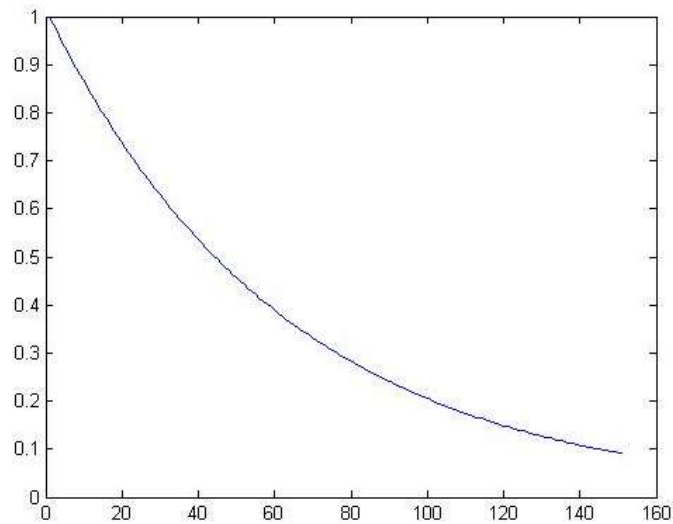
These pure signals are then mixed with different levels of random White Gaussian noise, with standard deviations  $\sigma = 0.00, 0.01, 0.05$ , respectively.

Noise  $\sim \mathcal{N}(0, 0.00); \mathcal{N}(0, 0.01); \mathcal{N}(0, 0.05)$ .

The second mono-exponential dataset,  $D_{mono}$ , is simulated in the same way as  $D_1$ . The pure signal decays are calculated using the mono-exponential model equation:

$$\text{Signal Decay} = e^{-bD_{mono}}$$

Noise simulation is done in the same way as for the bi-exponential model.



**Figure 15. Mono-exponential Signal:  $e^{-bD_{mono}}$**

The simulated data is a 2000-pixel image, with upper half containing bi-exponential data (Figure 14) and lower half containing mono-exponential data (Figure 15). Figure 16 shows the segmentation of the simulated data with bi-exponential region in black and mono-exponential region in white. Using the model selection process, the image is clearly segmented into two compartments (Section 6.2 – Figure 19).



**Figure 16. Distribution of Original Simulated Data**

## 5 Cramer-Rao Lower Bound (CRLB) for $D_{vas}$

---

### 5.1 Introduction to CRLB

The Cramer-Rao Lower Bound is the lowest possible bound on the variance of an unbiased estimator. In some cases, there are no existing estimators that can achieve that lower bound. If an unbiased estimator achieves it, the estimator is said to be efficient and is called the Minimum Variance Unbiased Estimator (MVUE).

CRLB is the inverse of the Fisher Information of the estimation parameter. Hence, the larger the Fisher Information value, the lower is the bound of the variance, i.e. the estimation is more accurate. Thus CRLB is used as a benchmark for comparison of performances of unbiased estimation techniques.

CRLB has been successfully used for various estimation problems like signal DC level estimation in WGN; phase and frequency estimation, in case of sinusoidal signals; estimation of parameters for line fitting, etc. For example, if we assume that the water fraction parameter  $f_{wat}$ , in our model, is a simple DC signal with WGN. The CRLB of  $f_{wat}$  can be calculated as follows:

$$f[n] = f_{wat} + w[n] \quad n = 0, 1, \dots, N-1, w[n] = \mathcal{N}(0, \sigma^2)$$

Likelihood of  $\mathbf{f}$ : 
$$p(\mathbf{f}; f_{wat}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (f[n] - f_{wat})^2 \right]$$

Log likelihood: 
$$\ln p(\mathbf{f}; f_{wat}) = -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (f[n] - f_{wat})^2$$

First derivative:

$$\begin{aligned} \frac{\partial \ln p(\mathbf{f}; f_{wat})}{\partial f_{wat}} &= \frac{\partial}{\partial f_{wat}} \left[ -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (f[n] - f_{wat})^2 \right] \\ &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (f[n] - f_{wat}) = \frac{N}{\sigma^2} (\bar{f} - f_{wat}) \end{aligned}$$

where  $\bar{f}$  is sample mean



Second derivative: 
$$\frac{\partial^2 \ln p(\mathbf{f}; f_{wat})}{\partial f_{wat}^2} = -\frac{N}{\sigma^2}$$

Fisher information: 
$$I(f_{wat}) = -E \left\{ \frac{\partial^2 \ln p(\mathbf{f}; f_{wat})}{\partial f_{wat}^2} \right\} = -E \left\{ -\frac{N}{\sigma^2} \right\} = \frac{N}{\sigma^2}$$

CRLB: 
$$\text{var}(f_{wat}) \geq \frac{1}{I(f_{wat})} = \frac{\sigma^2}{N}$$

The above result shows that the variance of  $f_{wat}$  can be estimated as low as the average of noise variance  $\left(\frac{\sigma^2}{N}\right)$ . The bound will decrease as sample size  $N$  increases or noise variance  $\sigma^2$  decreases.

However, in more general cases the calculation of CRLB is not always possible, e.g. when the first derivative of the log likelihood function does not exist at certain points of the interval.

## 5.2 Relevance of CRLB for the Estimation Methods Used

As discussed in Section 3.2, the proposed solution uses an iterative exhaustive search method to decrease the MSE of the parameters and applies both mono- and bi-exponential models on each voxel, to decide on the correct model.

Another common method of estimation is the least-squares curve fitting (MATLAB function – "*lsqnonlin*" ), which is used later in Section 6.4 for performance comparison with the proposed method.

Both these methods are sensitive to higher noise and statistical sampling errors. The proposed solution, using exhaustive search technique, can be deemed as optimal. However, to conclude that this solution is the best available, we need to verify that it gives better performance than the standard curve fitting technique and also gives comparatively closer values to a performance benchmark, like CRLB.

Since, the MSE of estimation is directly related to the estimated variance of the underlying parameters, a good way to judge the performance of these methods would be to compare the estimated variance with the CRLB value. Since, CRLB puts a lower limit on the variance (and resulting error), it can be used as a benchmark in this case. If the variance, from the proposed method, comes close to the CRLB, the method would be the best solution for this problem. In case, the estimation shows a wide difference from the CRLB, we can assume that better solutions might exist.

### 5.3 Derivation of CRLB for $D_{vas}$

Out of the three parameters in the bi-exponential model ( $f_{wat}, D_{wat}, D_{vas}$ ), described in Section 3.2.3, our interest mainly lies in the accurate estimation of the vascular diffusion coefficient –  $D_{vas}$ . Since vascular blood flow directly correlates to functional activity in the brain, the estimation of this parameter would help our understanding of the corresponding regions in the MRI images.

**From the bi-exponential model equation:**

$$y_n = f_{wat}(1 - b_n D_1) + (1 - f_{wat}) \exp(-b_n D_2) + w[n],$$

where  $n = \overline{0:N-1}$  and WGN  $w[n] = \mathcal{N}(0, \sigma^2)$ .

Let  $g_n(f, D_1, D_2) = f(1 - b_n D_1) + (1 - f) \exp(-b_n D_2)$

$$\therefore y_n = g_n(f, D_1, D_2) + w[n]$$

**Likelihood of y:**

$$p(\mathbf{y}; \mathbf{g}_n(f, \mathbf{D}_1, \mathbf{D}_2)) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{1}{2\sigma^2} (y_n - g_n(f, D_1, D_2))^2 \right]$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2))^2 \right]$$

**Log likelihood of  $y$ :**

$$\ln p(y; g_n(f, D_1, D_2)) = -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2))^2$$

$D_2$  can be estimated from the above log likelihood expression.

**First derivatives:**

$$\begin{aligned} \frac{\partial \ln p(y; g_n(f, D_1, D_2))}{\partial f} &= \frac{\partial}{\partial f} \left[ -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2))^2 \right] \\ &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial f} \end{aligned}$$

$$\begin{aligned} \frac{\partial \ln p(y; g_n(f, D_1, D_2))}{\partial D_1} &= \frac{\partial}{\partial D_1} \left[ -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2))^2 \right] \\ &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \end{aligned}$$

$$\begin{aligned} \frac{\partial \ln p(y; g_n(f, D_1, D_2))}{\partial D_2} &= \frac{\partial}{\partial D_2} \left[ -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2))^2 \right] \\ &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \end{aligned}$$

**Second derivatives:**

$$\begin{aligned} \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f^2} &= \frac{\partial}{\partial f} \left[ \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial f} \right] \\ &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left[ (y_n - g_n(f, D_1, D_2)) \frac{\partial^2 g_n(f, D_1, D_2)}{\partial f^2} - \left( \frac{\partial g_n(f, D_1, D_2)}{\partial f} \right)^2 \right] \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ln p(\mathbf{y}; \mathbf{g}_n(f, \mathbf{D}_1, \mathbf{D}_2))}{\partial \mathbf{D}_1^2} &= \frac{\partial}{\partial D_1} \left[ \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \right] \\
&= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left[ (y_n - g_n(f, D_1, D_2)) \frac{\partial^2 g_n(f, D_1, D_2)}{\partial D_1^2} - \left( \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \right)^2 \right] \\
\frac{\partial^2 \ln p(\mathbf{y}; \mathbf{g}_n(f, \mathbf{D}_1, \mathbf{D}_2))}{\partial \mathbf{D}_2^2} &= \frac{\partial}{\partial D_2} \left[ \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right] \\
&= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left[ (y_n - g_n(f, D_1, D_2)) \frac{\partial^2 g_n(f, D_1, D_2)}{\partial D_2^2} - \left( \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right)^2 \right] \\
\frac{\partial^2 \ln p(\mathbf{y}; \mathbf{g}_n(f, \mathbf{D}_1, \mathbf{D}_2))}{\partial f \partial \mathbf{D}_1} &= \frac{\partial}{\partial D_1} \left[ \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial f} \right] \\
&= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left[ (y_n - g_n(f, D_1, D_2)) \frac{\partial^2 g_n(f, D_1, D_2)}{\partial f \partial D_1} - \frac{\partial g_n(f, D_1, D_2)}{\partial f} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \right] \\
\frac{\partial^2 \ln p(\mathbf{y}; \mathbf{g}_n(f, \mathbf{D}_1, \mathbf{D}_2))}{\partial f \partial \mathbf{D}_2} &= \frac{\partial}{\partial D_2} \left[ \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial f} \right] \\
&= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left[ (y_n - g_n(f, D_1, D_2)) \frac{\partial^2 g_n(f, D_1, D_2)}{\partial f \partial D_2} - \frac{\partial g_n(f, D_1, D_2)}{\partial f} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right] \\
\frac{\partial^2 \ln p(\mathbf{y}; \mathbf{g}_n(f, \mathbf{D}_1, \mathbf{D}_2))}{\partial \mathbf{D}_1 \partial \mathbf{D}_2} &= \frac{\partial}{\partial D_2} \left[ \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (y_n - g_n(f, D_1, D_2)) \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \right] \\
&= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left[ (y_n - g_n(f, D_1, D_2)) \frac{\partial^2 g_n(f, D_1, D_2)}{\partial D_1 \partial D_2} - \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right]
\end{aligned}$$

**Fisher information:**

$$I(f, D_1, D_2) = -E \left[ \begin{array}{ccc} \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f^2} \right\} & \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_1} \right\} & \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_2} \right\} \\ \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_1} \right\} & \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_1^2} \right\} & \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_1 \partial D_2} \right\} \\ \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_2} \right\} & \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_1 \partial D_2} \right\} & \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_2^2} \right\} \end{array} \right]$$

$$= \begin{bmatrix} -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f^2} \right\} & -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_1} \right\} & -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_2} \right\} \\ -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_1} \right\} & -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_1^2} \right\} & -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_1 \partial D_2} \right\} \\ -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial f \partial D_2} \right\} & -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_1 \partial D_2} \right\} & -E \left\{ \frac{\partial^2 \ln p(y; g_n(f, D_1, D_2))}{\partial D_2^2} \right\} \end{bmatrix}$$

$$= \frac{1}{\sigma^2} \begin{bmatrix} E \left\{ \left( \frac{\partial g_n(f, D_1, D_2)}{\partial f} \right)^2 \right\} & E \left\{ \frac{\partial g_n(f, D_1, D_2)}{\partial f} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \right\} & E \left\{ \frac{\partial g_n(f, D_1, D_2)}{\partial f} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right\} \\ E \left\{ \frac{\partial g_n(f, D_1, D_2)}{\partial f} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \right\} & E \left\{ \left( \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \right)^2 \right\} & E \left\{ \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right\} \\ E \left\{ \frac{\partial g_n(f, D_1, D_2)}{\partial f} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right\} & E \left\{ \frac{\partial g_n(f, D_1, D_2)}{\partial D_1} \cdot \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right\} & E \left\{ \left( \frac{\partial g_n(f, D_1, D_2)}{\partial D_2} \right)^2 \right\} \end{bmatrix}$$

Expected value of the  $(y_n - g_n(f, D_1, D_2))$  term in the summation is zero.

Substituting first derivatives of  $g_n(f, D_1, D_2)$ :

$$\frac{\partial g_n(f, D_1, D_2)}{\partial f} = 1 - b_n D_1 - \exp(-b_n D_2),$$

$$\frac{\partial g_n(f, D_1, D_2)}{\partial D_1} = -f b_n,$$

$$\frac{\partial g_n(f, D_1, D_2)}{\partial D_2} = -b_n (1 - f) \exp(-b_n D_2)$$

$\therefore I(f, D_1, D_2)$

$$= \frac{1}{\sigma^2} \begin{bmatrix} \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2))^2 & -f \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2)) b_n & -f \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2)) b_n \exp(-b_n D_2) \\ -f \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2)) b_n & f^2 \sum_{n=0}^{N-1} b_n^2 & f(1 - f) \sum_{n=0}^{N-1} b_n^2 \exp(-b_n D_2) \\ -(1 - f) \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2)) b_n \exp(-b_n D_2) & f(1 - f) \sum_{n=0}^{N-1} b_n^2 \exp(-b_n D_2) & (1 - f)^2 \sum_{n=0}^{N-1} b_n^2 \exp(-2b_n D_2) \end{bmatrix}$$

**CRLB:**

$$[\text{cov}(\mathbf{f}, \mathbf{D}_1, \mathbf{D}_2)] \geq [\mathbf{I}(\mathbf{f}, \mathbf{D}_1, \mathbf{D}_2)]^{-1}$$

$$\text{Let } \mathbf{I}(\mathbf{f}, \mathbf{D}_1, \mathbf{D}_2) = \frac{1}{\sigma^2} \begin{bmatrix} A & B & C \\ B & D & E \\ C & E & F \end{bmatrix}$$

$$\Rightarrow [\mathbf{I}(\mathbf{f}, \mathbf{D}_1, \mathbf{D}_2)]^{-1} = \sigma^2 \begin{bmatrix} A & B & C \\ B & D & E \\ C & E & F \end{bmatrix}^{-1}$$

$$\Rightarrow \det = ADF + 2BCE - AE^2 - B^2F - C^2D$$

$$\Rightarrow \text{CRLB}(\mathbf{D}_2) = \frac{\sigma^2 \cdot \text{cofactor of } F}{\det} = \frac{\sigma^2 \cdot (AD - B^2)}{ADF + 2BCE - AE^2 - B^2F - C^2D}$$

Let  $\mathbf{b}_n = \mathbf{c}n$ , where  $c$  is the sampling interval,  
and assuming  $N$  is very large

$$A = \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2))^2 \cong \frac{c^2 N^3 D_1^2}{3} - \frac{3}{2c D_2}$$

$$B = -f \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2)) b_n \cong f \left[ \frac{c^2 N^3 D_1}{3} + \frac{1}{c D_2^2} \right]$$

$$C = -(1-f) \sum_{n=0}^{N-1} (1 - b_n D_1 - \exp(-b_n D_2)) b_n \exp(-b_n D_2) \cong (1-f) \left[ \frac{4k-3}{4c D_2^2} \right]$$

where  $k = \frac{D_1}{D_2}$  is a constant

$$D = f^2 \sum_{n=0}^{N-1} b_n^2 \cong \frac{c^2 N^3 f^2}{3}$$

$$E = f(1-f) \sum_{n=0}^{N-1} b_n^2 \exp(-b_n D_2) \cong \frac{f(1-f)}{c D_2^3}$$

$$F = (1-f)^2 \sum_{n=0}^{N-1} b_n^2 \exp(-2b_n D_2) \cong \frac{(1-f)^2}{4c D_2^3}$$

$$\begin{aligned} \text{As } \text{CRLB}(D_2) &= \frac{\sigma^2 \cdot (AD - B^2)}{ADF + 2BCE - AE^2 - B^2F - C^2D} \\ &\cong \sigma^2 \cdot \frac{8(4k + 3)cD_2^3}{(1 - f)^2(8k + 15)} \end{aligned}$$

$$\therefore \text{CRLB}(D_2) \cong \frac{Kc\sigma^2 D_2^3}{(1 - f)^2}, \quad \text{where } K = \frac{8(4k + 3)}{(8k + 15)} \text{ is a constant}$$

**Final expression for CRLB:**

$$\text{var}(D_{vas}) \geq \frac{KcD_{vas}^3\sigma^2}{(1 - f_{wat})^2}$$

## 5.4 Observations from the Derived CRLB

From the above expression it can be seen that variance of  $D_{vas}$  depends on the following:

- Noise variance ( $\sigma^2$ ) – As noise variance increases, the estimation becomes worse.
- Water volume fraction ( $f_{wat}$ ) – As water fraction approaches 1, the expression approaches to  $\infty$ , showing that the estimation becomes impossible.
- Vascular diffusion coefficient ( $D_{vas}$ ) – As the CRLB expression is directly proportional to  $D_{vas}^3$ , even a small increase in  $D_{vas}$  results in a significant increase in the CRLB, making the estimation unusable.
- Sampling interval  $c$  – This is a user controlled constant value. The CRLB is directly proportional to this constant and an increase in the value will increase CRLB as well, making the overall performance worse. However, if  $D_{vas}^3$  becomes large, the value of  $c$  can be reduced (and the sample size  $N$  correspondingly increased), to keep the term  $cD_{vas}^3$  constant. This will result in more number of b-values.

Since, the numerator term  $cD_{vas}^3$  can be kept constant by controlling  $c$ , the CRLB expression can be reduced to a function of the remaining two parameters –  $\sigma^2$  and  $f_{wat}$ .

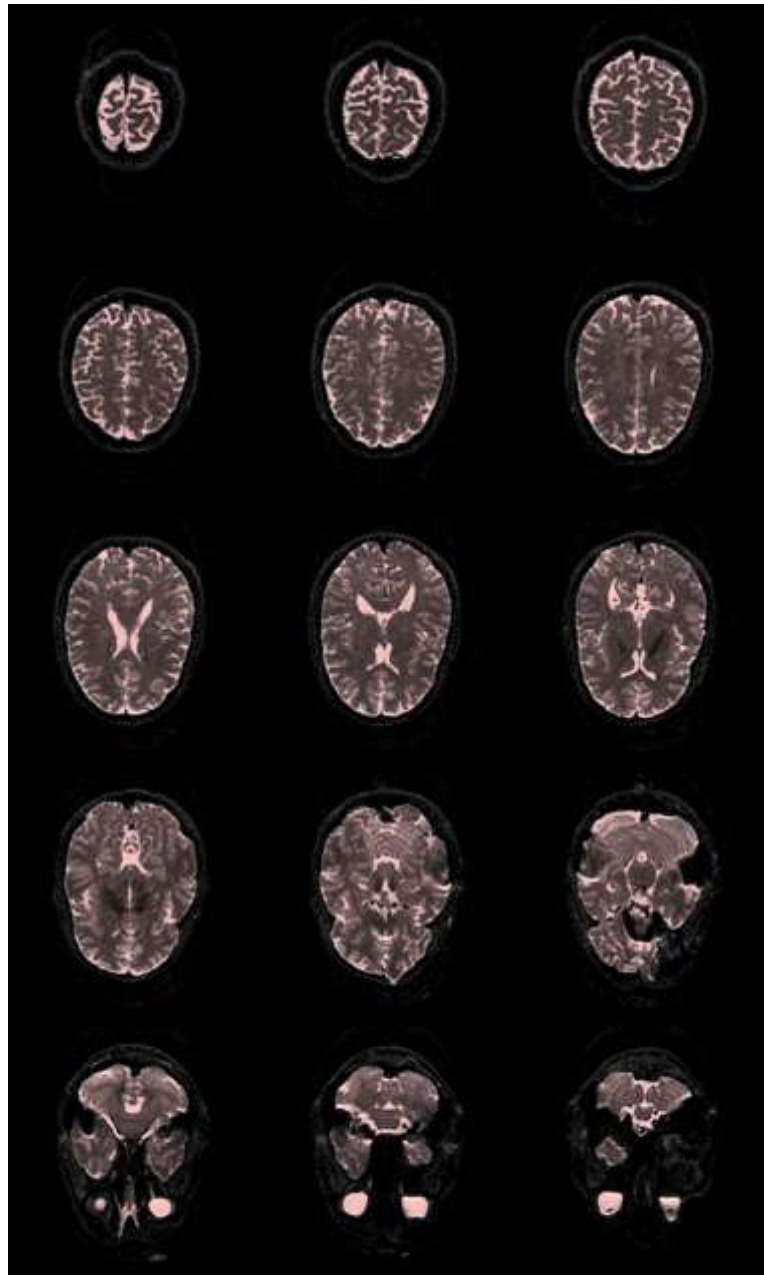


## 6 Experiment Results and Discussion

---

### 6.1 Data Pre-processing Results

Figure 17 below was generated by MRicro. The red pixels are the areas located within the boundary. The areas outside the boundary are the gray pixels (skull) and the background.



**Figure 17. Skull Removal Result**

A threshold value of 900 is chosen as the voxel intensity to identify the background and the skull voxels. All voxel intensities below the threshold are defined as the background and the skull. The other voxels are considered as the regions of interest for our experiment.

## 6.2 Model Selection Results for Simulated Data

To determine the level of accuracy of the model selection technique, we process the simulated data first, for different values of the parameters –  $f_{wat}$  and  $\sigma$ . The reason for choosing these two parameters for testing is the fact that the CRLB expression for  $D_{vas}$  directly depends on them (Section 5.4). Hence, the results can be used to set up a comparison between the estimated variance values of  $D_{vas}$  and the CRLB values for the various levels of  $f_{wat}$  and  $\sigma$ .

**Table 7. Estimation of  $D_{vas}$  subject to Noise Level and Water Fraction**

Noise	$f_{wat}$	$Sqrt(MSE(\hat{D}_{vas}))/mean(D_{vas})$
No noise	0.80	0.0000
No noise	0.85	0.0000
No noise	0.90	0.0000
$\mathcal{N}(0,0.01)$	0.80	0.0515
$\mathcal{N}(0,0.01)$	0.85	0.0735
$\mathcal{N}(0,0.01)$	0.90	0.1040
$\mathcal{N}(0,0.05)$	0.80	0.2711
$\mathcal{N}(0,0.05)$	0.85	0.4288
$\mathcal{N}(0,0.05)$	0.90	0.9114

Table 7 shows the ratio of the root mean squared errors over the mean value of  $D_{vas}$ . These ratios show how erroneous the estimated value is, compared to its true value.

The Mean Squared Error (MSE) for the estimator  $D_{vas}$ , is computed for each dataset using the following formula:

$MSE(\hat{D}_{vas}) = [\sum(\hat{D}_{vas} - D_{vas0})^2]/N$ , where  $\hat{D}_{vas}$  and  $D_{vas0}$  are estimated and original values of  $D_{vas}$ , respectively; and N is the number of pixels.

The 3-D graph in Figure 18 illustrates the values given in Table 7. It can be observed that error levels of  $D_2$  increase steadily with higher noise and higher values of  $f_{wat}$ .

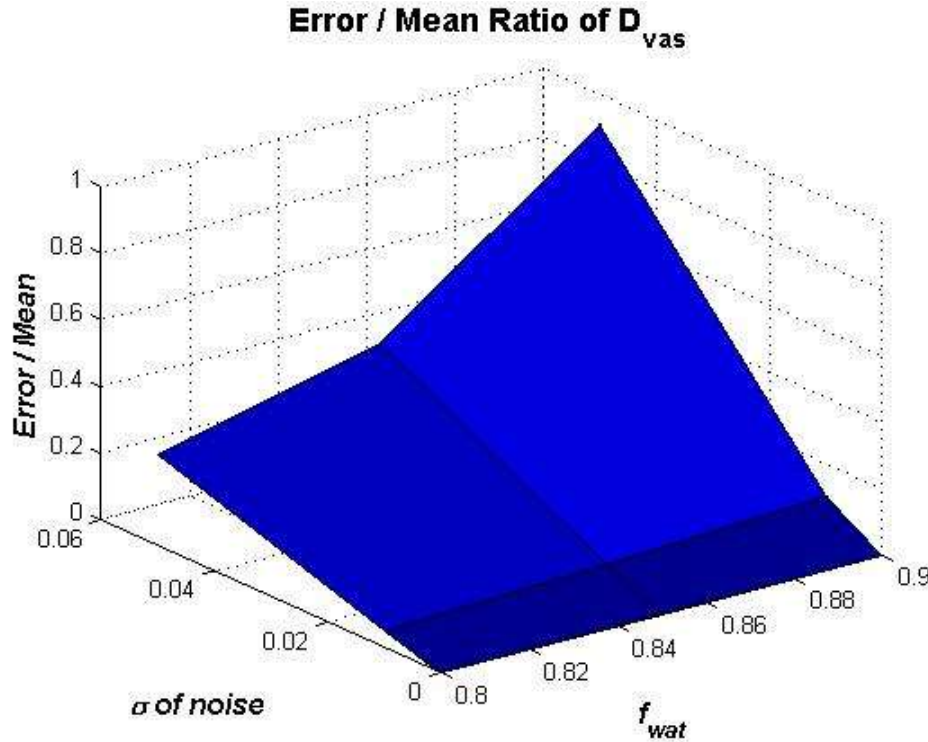


Figure 18. Error Levels of  $D_{vas}$

The images in Figure 19 show the estimated separation of the simulated data with and without application of random Gaussian noise  $\sigma \sim \mathcal{N}(0, 0.05)$ , at  $f_{wat} = 0.8$ .

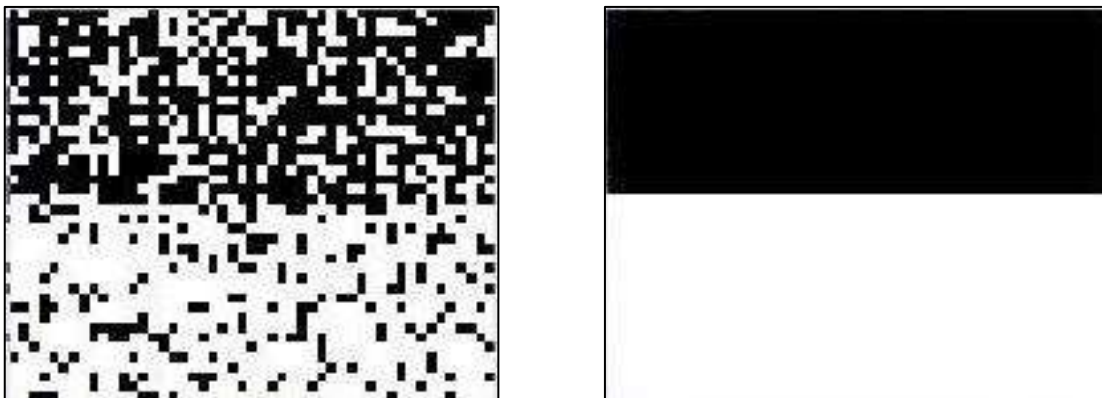


Figure 19. Estimated Separation of Both Models with and without Noise

This is done using the model selection process on each pixel in the image. The result in Figure 19 clearly shows the separation between two sets of data (two halves of the image), with mono-exponential pixels shown in white and bi-exponential pixels shown in black colour.

The ratio of  $\text{Sqrt}(MSE(\hat{D}_{vas}))/\text{mean}(D_{vas})$  at  $\sigma \sim \mathcal{N}(0,0.05)$ ,  $f_{wat} = 0.8$  is observed to be approximately 0.2711 (ref. Table 7). This error can be seen in the estimated separation in Figure 19 (left) above.

### 6.3 Comparison with CRLB Values

$$\text{CRLB} = \frac{KcD_{vas}^3\sigma^2}{(1 - f_{wat})^2}$$

Parameter values for the CRLB expression can be from Section 4.3 (data simulation):

- $N = 151$ .
- $c = 16 \Rightarrow b_n = 16n = 0,16, \dots 2400 \text{ s/mm}^2$ .
- $D_{vas} = 7 \times 10^{-3} \text{ mm}^2/\text{s}$ .
- WGN  $\sigma = 0, 0.01, 0.05$ .
- $f_{wat} = 0.80, 0.85, 0.90$ .

**Table 8. Estimated Error vs. CRLB for Different Levels of  $\sigma$  and  $f_{wat}$**

$\sigma$	$f_{wat}$	$\text{sqrt}(\text{CRLB}) \text{ } 10^{-3} \text{ mm}^2/\text{s}$	$\text{sqrt}(\text{Err. of } D_{vas}) \text{ } 10^{-3} \text{ mm}^2/\text{s}$
0.00	0.80	0.0000	0.0000
0.00	0.85	0.0000	0.0000
0.00	0.90	0.0000	0.0000
0.01	0.80	0.3202	0.3606
0.01	0.85	0.4269	0.5148
0.01	0.90	0.6404	0.7280
0.05	0.80	1.6009	1.8974
0.05	0.85	2.1346	3.0017
0.05	0.90	3.2019	6.3796

From Table 8 – CRLB values, we observe that as the noise levels or water fraction increase, the CRLB value becomes higher. This behaviour is as expected. Also, the differences between the lower bound value and the estimated values are very small, which shows the effectiveness of the proposed estimation model.

## 6.4 Comparison with MATLAB “lsqnonlin” Function

The method could be compared with another method that makes use of the MATLAB non-linear fitting function "*lsqnonlin*". This MATLAB function requires initialisation of all the parameters and hence different initial values give a variation of results.

**Table 9. Comparison of Results with MATLAB Non-linear Fitting Function**

$\sigma$	$f_{wat}$	$\text{sqrt}(CRLB)$ $10^{-3} \text{ mm}^2/s$	$\text{sqrt}(Err. \text{ of } D_{vas})$ $10^{-3} \text{ mm}^2/s$ with proposed method	$\text{sqrt}(Err. \text{ of } D_{vas})$ $10^{-3} \text{ mm}^2/s$ with lsqnonlin
0.00	0.80	0.0000	0.0000	0.0000
0.00	0.85	0.0000	0.0000	0.0000
0.00	0.90	0.0000	0.0000	0.0000
0.01	0.80	0.3202	0.3606	4.9475
0.01	0.85	0.4269	0.5148	5.0709
0.01	0.90	0.6404	0.7280	5.1705
0.05	0.80	1.6009	1.8974	5.4473
0.05	0.85	2.1346	3.0017	5.9549
0.05	0.90	3.2019	6.3796	6.0740

Table 9 depicts the results obtained from the MATLAB function, with initial values set as the real values of the three parameters. The results are compared against the CRLB levels, as well as the values obtained from the proposed method in Table 8.

It can be clearly observed from Table 9, that compared to both CRLB values and the results from the proposed method, the error levels from the “lsqnonlin” function are

much worse, except for the last point, where  $\sigma$  and  $f_{wat}$ , both become high and the error is slightly higher (as expected from Section 5.4).

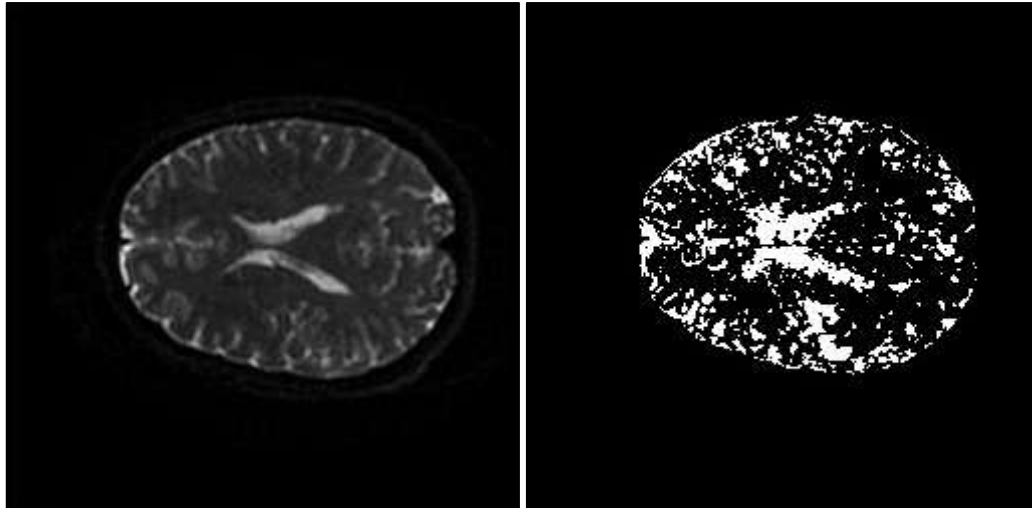
From the above results; and keeping in mind the fact that “lsqnonlin” function required all three parameters to be initialised to their true values for the above estimation, whereas the proposed solution only initialised  $D_{vas}$  at a default value of 0, we can conclude that our method using the Exhaustive Search strategy has a better performance than the MATLAB nonlinear fitting function, which is commonly used for solving problems of this category.

## 6.5 Model Fitting Results on Real MRI Signal Data

The next step is to test the procedure on real MRI signal data. The data is fitted into both mono-exponential and bi-exponential models, and then the model selection process is applied. The results provide the estimated number of compartments in each of the brain regions in the image. These results can be compared with ground truths and previous results to evaluate the consistency and efficiency of the process.

As far as the distribution of the number of compartments is concerned, it is known that the GM and WM regions are better fitted with the bi-exponential model, because capillaries with blood flow are more prominent in these regions. However, the CSF being a fluid, the mono-exponential model fits these voxels best, because this region mainly contains isotropic water diffusion.

Figure 20 shows the result of the model selection process on a sample MRI image. The white voxels depict mono-exponential one-compartment CSF areas, whereas the black voxels depict the bi-exponential two-compartment WM and GM areas. The distribution of these voxels is quite reasonable as observed in the resultant image.



**Figure 20. Original MRI Image vs. Number of Compartments Distribution**

But, from observation of the residual sum of square errors from the fitting of both models, given in Table 10, it can be seen that CSF voxels have a higher error rate than WM and GM voxels. The root MSE rate for the mono-exponential model in the CSF region is more than 2 times higher than the other regions. This is against the assumption that the mono-exponential model would fit the CSF region better than the WM and GM regions.

**Table 10. Root Mean Squared Errors of CSF vs. WM & GM**

	CSF ( $10^{-3} \text{ mm}^2/\text{s}$ )	WM & GM ( $10^{-3} \text{ mm}^2/\text{s}$ )
Mono-exponential Model	3.7417	1.7321
Bi-exponential Model	3.1623	1.7321

This anomaly can be explained with the fact that the CSF regions have a much higher noise level ( $\sigma > 0.1$ ). It has been shown earlier that with such high noise, the F-test for the model selection becomes unusable. In such case, the model selection threshold was amended from 0 to -0.002 to give acceptable results.

On comparing results between Table 8 (error on simulated data) and Table 10 (error on real data), we can deduce that for WM and GM, the error level corresponds to  $\sigma = 0.05, f = 0.80$ .



The next test is to measure the discrepancies between the parameter estimates from our method against their true values. Table 11 shows the estimated values of the various parameters involved.

**Table 11. Estimated parameters in Bi-exponential and Mono-exponential voxels**

$f_{wat}$ (Whole brain)	$f_{wat}$ (Two-comp region)	Noise (Two-comp region)	$D_{wat}$ ( $10^{-3}mm^2/s$ )	$D_{vas}$ ( $10^{-3}mm^2/s$ )	$D_{mono}$ ( $10^{-3}mm^2/s$ )
$0.86 \pm 0.15$	$0.8 \pm 0.14$	$\sim \mathcal{N}(0, 0.05)$	$0.76 \pm 0.02$	$7.11 \pm 5.31$	$1.76 \pm 0.37$

**Table 12. The Ground Truth**

Parameter	Value ( $10^{-3}mm^2/s$ )
Fast Diffusion coefficient(Extracellular Diffusion)[10]	$0.824 \pm 0.03$
Fast Diffusion fraction[10]	$0.8 \pm 0.02$
Human: Diffusion coefficient of tissue water	1.155 – 1.479
Cat: Diffusion coefficient of tissue water	0.8

Table 12 shows the true mean values of these parameters and also some statistical averages.

Comparing the data from both the tables shows that the estimated values are very close to the expected values. The estimates of the water diffusion coefficient and its volume fraction are identical to the true values.

Recalling the previous discussion on the accuracy of the proposed method with simulated data, the tolerance level of  $D_2$  was within a range of  $4 \times 10^{-3}mm^2/s$ , which enables us to detect significant changes of blood flow in any of the voxels. This proves as a very useful starting point for future studies, using the estimated values of the vascular diffusion coefficients at the individual voxel level.



## 7 Conclusions and Future Work

---

### 7.1 Conclusions

The experimental results are compared to the statistical data collected for brain images from cat brains. From the discussions and experimental results, on both simulated and real MRI data, it is clear that the proposed solution gives highly acceptable results. Also, from the results in Table 10 and Table 11, it can be concluded that the assumption of WGN distribution, for the noise levels in real MRI data, is valid.

The model selection method proposed is based on an F-test, by applying a threshold of zero on the weighted difference:  $(A \times RSS_1 - RSS_2)$ . If the difference is lesser than or equal to threshold value zero, mono-exponential model is selected, else bi-exponential model is selected. Here  $A$  is a constant calculated from the standard value of  $F$  (based on 5% significance level) and the degrees of freedom for both models.

However, as the noise levels go up ( $\sigma > 0.1$ ), the performance of the technique goes down and makes it unusable for any practical purposes. This is one of the major constraints for the proposed technique, at the moment.

Another constraint to the proposed technique is the usage of only first partial derivatives of parameters for minimisation of RSS value, and not considering the second partial derivatives, due to the high sensitivity to noise for these derivatives.

The proposed techniques of Exhaustive Search for initial parameter estimation and the actual model fitting and selection process to distinguish individual voxels into mono-exponential and bi-exponential models, proves to give estimates very near to the true mean values.

Comparison of the estimate to the CRLB value obtained for the variance of vascular diffusion coefficient ( $D_{vas}$ ) also show high proximity, further confirming the efficacy of this model.

## **7.2 Future Work**

One of the major challenges of the model selection process is the high noise level in MRI images. Above a certain limit, the noise makes the process almost ineffective in determining the correct model. Also, the calculation of second partial derivatives for minimising RSS gets affected by the increase in noise level. A possible solution for this problem could be pre-estimation of the actual noise distribution in the image, so that the noise can be eliminated as far as possible and results can be obtained more accurately.

Another challenge is to utilise neighbourhood voxel information in the voxel-by-voxel estimations given by this method. As of now, the method disregards the neighbouring voxels for the individual calculations. The fact that a strong relationship exists among the diffusion levels of neighbouring voxels from the same region, cannot be ignored. As of now, the method only detects water diffusion and vascular diffusion compartments. The inclusion of neighbourhood information in the calculations may possibly enhance the method to help to distinguish between the different regions (GM, WM and CSF).

Based on the Fisher Information and CRLB values obtained in Chapter 5, another possible area of study for the future would be to improve on the estimation methods further, in order to more closely achieve the CRLB, to improve the accuracy of the solution.

Enhancements to our method in these three areas are plausible candidates for future research in this subject. These improvements would help to make the method more robust and efficient.

## References

---

- [1] D. Le Bihan, J. Mangin, and C. Poupon *et al.*, "Diffusion tensor imaging: concepts and applications," *Journal of Magnetic Resonance Imaging*, vol. 13, no. 4, pp. 534–546, 2001.
- [2] R. A. Pooley, "AAPM/RSNA Physics Tutorial for Residents: fundamental physics of MR imaging," *Radiographics*, vol. 25, no. 4, pp. 1087–1099, 2005.
- [3] P. J. Basser, and D.K. Jones, "Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review," *NMR Biomed*, vol. 15, no. 7–8, pp. 456–467, 2002.
- [4] S. Mori, *Introduction to Diffusion Tensor Imaging*. Elsevier, 2007.
- [5] S. Chabert, and P. Scifo, "Diffusion signal in magnetic resonance imaging: origin and interpretation in neurosciences," *Biol. Res.*, vol. 40, no. 4, pp. 385–400, 2007.
- [6] S. C. Steenset *al.*, "Reproducibility of brain ADC histograms," *Eur Radiol*, vol. 14, no. 3, pp. 425–430, 2004.
- [7] T. Jin, F. Zhao, and S.G. Kim, "Sources of functional apparent diffusion coefficient changes investigated by diffusion-weighted spin-echo fMRI," *Magn. Reson. Med.*, vol. 56, no. 6, pp. 1283–1292, 2006.
- [8] E. Salaet *al.*, "Apparent diffusion coefficient and vascular signal fraction measurements with magnetic resonance imaging: feasibility in metastatic ovarian cancer at 3 Tesla: technical development," *Eur Radiol*, vol. 20, no. 2, pp. 491–496, 2010.

- [9] H. J. Motulsky, and L.A. Ransnas, "Fitting curves to data using nonlinear regression: a practical and nonmathematical review," *FASEB J.*, vol. 1, no. 5, pp. 365–374, 1987.
  
- [10] T. Niendorf *et al.*, "Biexponential diffusion attenuation in various states of brain tissue: implications for diffusion-weighted imaging," *Magn. Reson. Med.*, vol. 36, no. 6, pp. 847–857, 1996.