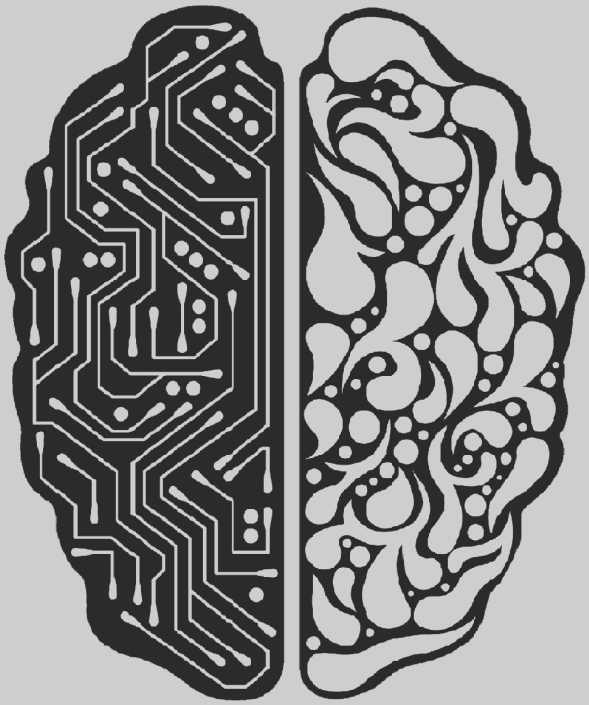
University of Crete
Mathematics and Applied Mathematics Department

Diffusion Magnetic Resonance Imaging Techniques: Applications in Brain and Human Body



Master Thesis

Georgios Ioannidis

Supervisor: Kampanis Nikos *, Ph.D., Research Director
Evaluators: Marias Kostas**, Ph.D., Principal Researcher
Tsogka Chrysoula***, Professor

* Institute of Applied and Computational Mathematics, FORTH
** Institute of Computer Science, FORTH
*** Department of Applied Mathematics, University of Crete

Abstract

Medical imaging is the process of creating visual representations of the interior of our body for clinical analysis and medical intervention. Medical imaging reveals the internal structures hidden by the skin and it is a powerful tool for population screening and efficient patient management. Thirty years ago, Magnetic Resonance Imaging (MRI) was introduced in clinical practice and was progressively established among the most frequently used medical imaging modalities. Apart from conventional MR imaging, more advanced techniques were arisen such as spectroscopy, diffusion and perfusion MRI.

The main subject of this work is Diffusion Weighted Imaging (DWI) and is focused on obtaining diffusion parameters from the Intravoxel Incoherent Motion (IVIM) model which gives information about water diffusivity and tissue vascularity. In order to extract the aforementioned parameters from the IVIM model, non-linear least squares are used. As an iterative process, non-linear least squares need a starting point (initial guess), a lower and an upper bound. The initial point is of crucial significance for the convergence of the algorithm. Due to the complexity of the IVIM function (many local minima), standard non-linear least squares fail to produce clinically reliable results inside the natural bounds.

To this end, we present a novel mathematical framework which exploits the characteristics of the Diffusion Weighted (DW) signal curve in order to find the appropriate initial starting points and bounds for the non-linear least squares. Comparisons between results obtained by the proposed and those of the standard initialization, for the non-linear least squares, results in clinically relevant and reliable results with our method, in contrast to the standard initialization which suffers from reproducibility of the results.

Περίληψη

Ιατρική απεικόνιση είναι η διαδικασία της δημιουργίας οπτικής αναπαραστάσης του εσωτερικού του σώματος μας για κλινική ανάλυση και ιατρική παρέμβαση. Η ιατρική απεικόνιση αποκαλύπτει τις εσωτερικές δομές, κρυμμένες από το δέρμα και αποτελεί ισχυρό εργαλείο για τον έλεγχο του πληθυσμού και την αποτελεσματική διαχείριση των ασθενών. Τριάντα χρόνια πριν, η μαγνητική τομογραφία (MRI) εισήχθη στην κλινική πρακτική και σταδιακά καθιερώθηκε ως μία από τις πιο συχνά χρησιμοποιούμενες ιατρικές μεθόδους απεικόνισης. Εκτός από τη συμβατική απεικόνιση MR, έχουν προκύψει και άλλες προηγμένες τεχνικές, όπως η φασματοσκοπία, MRI διάχυσης και αιμάτωσης.

Το κύριο αντικείμενο της εργασίας αυτής είναι το MRI διάχυσης, και επικεντρώθηκε στην απόκτηση των παραμέτρων διάχυσης από το μοντέλο Intravoxel Incoherent Motion (IVIM), το οποίο παρέχει πληροφορία σχετικά με διαχυτότητα του νερού και την αγγείωση στους ιστούς. Προκειμένου να εξαχθούν οι προαναφερθείσες παράμετροι από το IVIM μοντέλο, χρησιμοποιούνται τα μη-γραμμικά ελάχιστα τετράγωνα. Ως επαναληπτική διαδικασία, τα μη-γραμμικά ελάχιστα τετράγωνα χρειάζονται ένα σημείο εκκίνησης, (αρχική εκτίμηση) ένα κάτω και ένα άνω φράγμα. Η αρχική συνθήκη είναι υψίστης σημασίας για τη σύγκλιση του αλγορίθμου. Λόγω της πολυπλοκότητας της συνάρτησης IVIM, (πολλά τοπικά ελάχιστα) τα μη-γραμμικά ελάχιστα τετράγωνα αποτυγχάνουν στο να παράγουν κλινικά αξιόπιστα αποτελέσματα εντός των φυσικών ορίων.

Με σκοπό αυτό, παρουσιάζουμε ένα νέο μαθηματικό πλαίσιο, το οποίο εκμεταλλεύεται τα χαρακτηριστικά της καμπύλης του σήματος διάχυσης, προκειμένου να βρεθούν οι κατάλληλες αρχικές συνθήκες και τα όρια για το αντίστροφο πρόβλημα. Συγκρίσεις μεταξύ των αποτελεσμάτων της προτεινόμενης και αυτών της καθιερωμένης αρχικοποίησης των μη-γραμμικών ελαχίστων τετραγώνων, οδηγούν σε κλινικά σημαντικά και αξιόπιστα αποτελέσματα με τη μέθοδο μας, σε αντίθεση με την καθιερωμένη αρχικοποίηση που υποφέρει από την επαναληψημότητα των αποτελεσμάτων.

Acknowledgements

I would like to express my special appreciation and thanks to the director of the Computational Bio-Medicine Laboratory (CBML) Dr. Kostas Marias for his valuable advice and guidance and for the excellent work environment. Without whom this dissertation would not have been possible.

Many thanks go to my supervisor Dr. Kampanis Nikos for his useful guidance.

I would like to express my gratitude to my colleagues Georgios C. Manikis, Katerina Nikiforaki and Lefteris Kontopodis for introducing me to the topic and for the useful comments and remarks.

I would also like to thank my committee members, Dr. Kampanis Nikos, Dr. Marias Kostas, and Prof. Tsogka Chrysoula for their involvement.

Last but not least, I would like to thank my family, my friends, my colleagues from CBML and of course my partner in life for their continuous support.

Contents

List of Figures

1

Introduction

Throughout the years we "struggled" to understand important organs in our body. In the last decades our kind has come to a new era in investigating the mysteries of our body with revolutionary techniques. The following chapters, are focused on Magnetic Resonance Imaging (MRI) and precisely in Diffusion Weighted MRI (DWI), a fast and non-invasive technique in which, no-contrast media or ionizing radiation are required.

DWI is a technique that can give qualitative information about the anatomy and it is possible to characterize tissues and diseases [2]. The Apparent Diffusion Coefficient. (ADC), is a significant bio-marker that is derived from the DWI data and it can provide a quantitative measure of water diffusivity, which mirrors tissue cellularity and tissue organization. Furthermore, DWI has been used for the assessment of abdominal malignancies and it is capable to characterize tumours in organs such as, the kidneys, the liver and the pancreas [3].

Abdominal lesions of liver or pancreas have high occurrence and thus the differentiation between benign or malignant nature of the lesion depends highly on MR findings. DWI techniques will be applied to the diffusion MRI data coming from regions of the liver and pancreas, thus a short description of their anatomy and structure is essential for this work.

In the next pages, the MRI principles and the basic diffusion concepts are described, along with a detailed description concerning the theoretical background of DWI. The Intravoxel Incoherent Motion (IVIM) model is applied in the abdominal area in order to give clinically parameters for diffusion. IVIM is a more advanced model that exploits the effect on the DW signal due to the micro-circulation of blood. These parameters are a result of non-linear regression. Therefore one of the following chapters will cover the mathematical background of the non-lineal least squares theory along with our modifications in order to provide optimized results from the DWI analysis.

1.1 The Human Brain

The brain is part of the central nervous system (C.N.S.) where all organs and properties of the human body are controlled, but also where perception of sensory input and higher conscious and mental processes take place. So, the brain is the root of what makes us human.

Brain tissue structure is formed by different types of neuronal and glial cells. Neurons are the most important components because of their ability of receiving and integrating information, data processing and transmission through their axonal projections and

synaptic junctions (the contact points between neurons) to other cells. A typical neuronal cell consists of: the cellular body, which is the main part of the cell, the dendrites, the signal receiver and the axon, the signal transmitter. A tract is a large number of afferent or efferent, myelinated or not, axons that defines a path and wires brain regions.

Axons, these cellular protrusions that help signal transduction between cells, are either sheathed with myelin or non-myelinated. Myelin layer enables more efficient signalling when needed in the neural network. Myelin is constructed by glial cells and insulation of electrical conductivity is one of their main roles in C.N.S. apart from maintaining brain homeostasis, mechanical and trophic support, protection from pathogens and others [4].

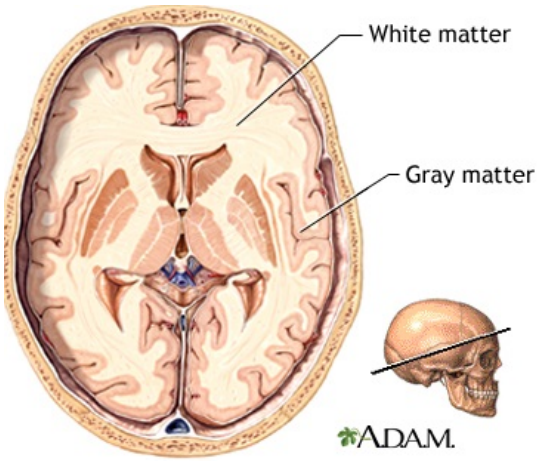


Figure 1.1: Gray and white matter of the brain

A worth mentioning property when brain slices are histologically described is that in regions rich in cellular bodies, the dominant color is grey, while white regions are those where myelinated axonal tracts are concentrated. Thus, brain tissue differentiates into interchangeable grey and white matter (figure 1.1). Regions between gray and white matter communicate via clusters of axons, called tracts. Those tracts are able to be depicted through imaging modalities, as presented in the next chapter.

1.1.1 Anatomy

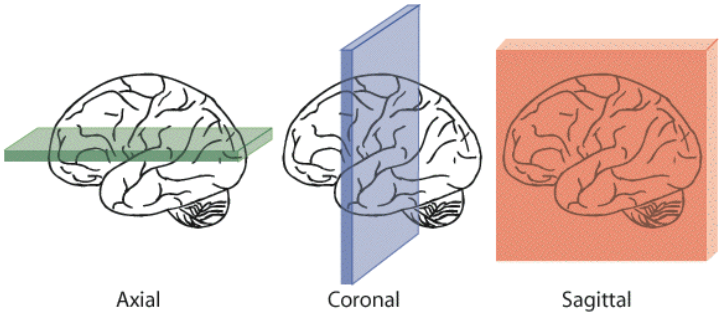


Figure 1.2: Planes of the brain. Figure taken by: pixgood.com

Macroscopically, the brain is not a homogeneously structured organ, but can be divided into three anatomical sections. Firstly, the cerebrum (forebrain) is composed by the cerebral hemispheres, which are furthermore divided into the four lobes (frontal, parietal, occipital, temporal), and also contains the diencephalon, where the thalamus lies in. Also, the cerebellum (hindbrain), which is also called the small brain. Finally, the brainstem which is subdivided into the midbrain, the pons and the medulla, which is

the bridge between the brain and the spinal cord. All these regions are functionally related to a uniform network, where every area is responsible for a set of functions. A brief generalized description is shown in (figure 1.3).

Driven by the three axes that may be established for the organ, the brain sections are three mutually perpendicular planes, either sagittal, axial (transverse) or coronal (figure 1.2). As it can be seen in the next pages, MR images follow the same rules.

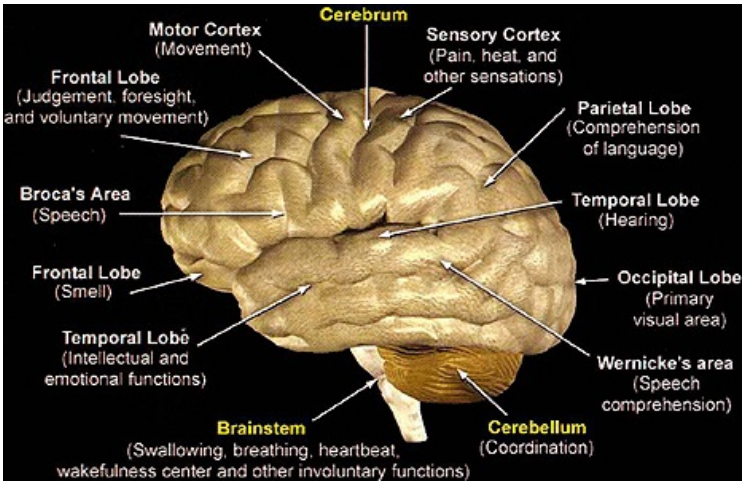


Figure 1.3: Gross anatomical brain structures. Figure taken by: University Hospital Newark, NJ.

1.1.2 Cerebrospinal fluid (CSF)

Within the brain mass, there is the ventricles' system (Figure 1.4) where the CSF is contained. The ventricles are spaces insulated from the brain tissue dedicated to the production of the CSF. The role of the CSF is twofold: it is involved in the metabolism of the cells of the C.N.S. and it functions as a protector of the brain from the external vibrations [4].

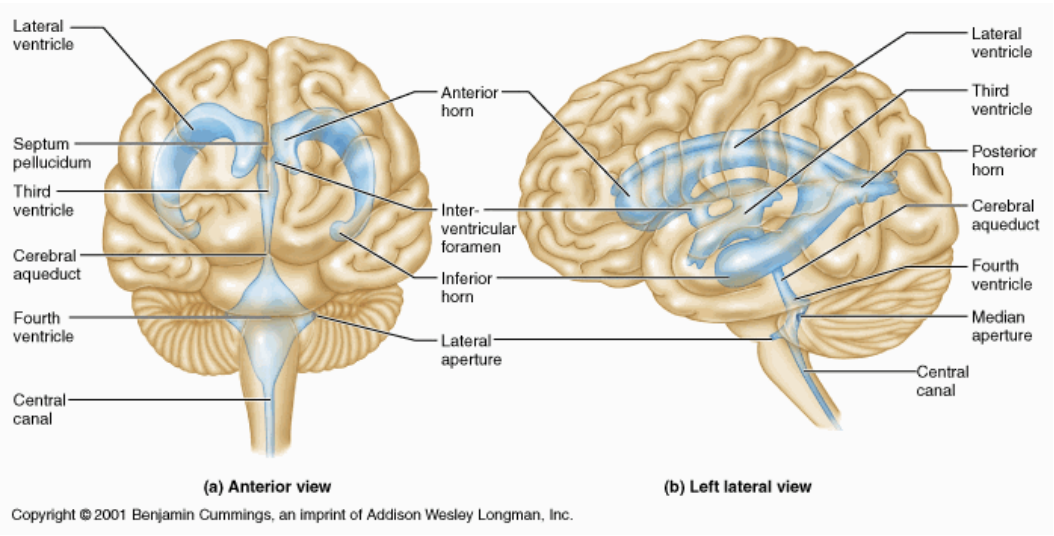


Figure 1.4: Ventricles' system.

The largest amount of water in the brain mass is located in the ventricles' system. However, CSF differs from blood in its content though, both of them obviously have water as the main component.

1.2 The Liver

The liver (figure 1.5) is a vital organ and it is located in the upper right quadrant of the abdomen, below the diaphragm. The liver has a wide range of functions, including detoxification of various metabolites, protein synthesis, and the production of biochemicals necessary for digestion. The liver is a gland and plays a major role in metabolism with numerous functions in the human body, including decomposition of red blood cells, plasma protein synthesis, and hormone production. Also it is highly vascularized organ with dense capillary system.

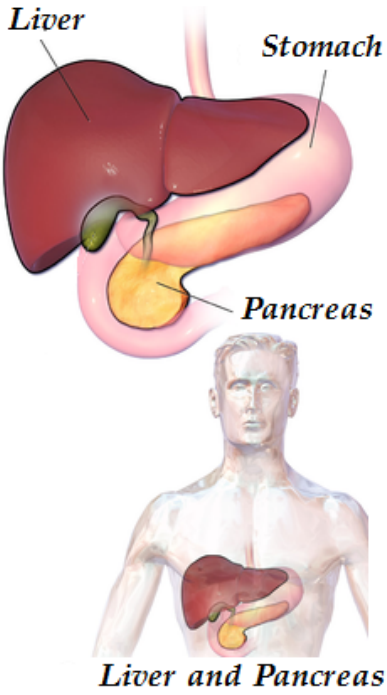


Figure 1.5: Abdominal organs. Liver and Pancreas

1.3 The Pancreas

The pancreas (figure 1.5) is a glandular organ in the digestive system and endocrine system of vertebrates. In humans, it is located in the abdominal cavity behind the stomach. It is an endocrine gland producing several important hormones, including insulin, glucagon, somatostatin, and pancreatic polypeptide which circulate in the blood. The pancreas is also a digestive organ, secreting pancreatic juice containing digestive enzymes that assist digestion and absorption of nutrients in the small intestine. These enzymes help to further break down the carbohydrates, proteins, and lipids in the chyme.

2

Magnetic Resonance Imaging

As pathology is not uncommon for the aforementioned specific organs, anatomical and functional imaging is of paramount importance. Apart from conventional, functional information, is necessary in order to aid diagnosis. MRI is the modality of choice, as multi parametric MR acquisition can give an insight into tissue integrity, cellularity and vascularization.

In order to analyse the water diffusion in the human body and brain through **M**agnetic **R**esonance **I**maging (**MRI**), we have to understand the basic physics behind **MRI**. To begin with, we have to determine the terms "z-direction" and "xy-plane" that are frequently used in this chapter. In all figures, the main magnetic field, B_0, is represented from bottom to top and its direction is designated by z. The other two dimensions of the magnetic field are denoted by x and y. The xy-plane is perpendicular to the z-axis and is represented horizontally in figure 2.1.

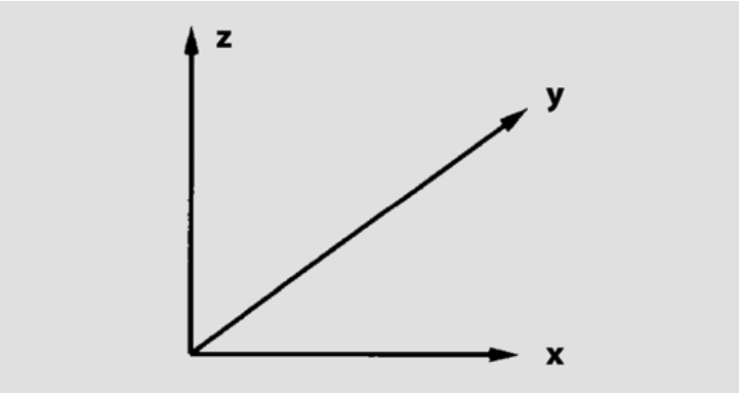


Figure 2.1: Magnetization Directions: z-axis ,xy-plane

2.1 Spin and the Nuclear Magnetic Resonance Phenomenon

The exact description of what happens when a magnetic field is applied on human tissue, relies on quantum mechanics. All the theory necessary for MRI can be based on a simple and precise model in which the protons can be considered as small magnets. MRI uses the signal from the nuclei of hydrogen atoms (^1H) to generate an image, utilizing the fact that living tissue consists of 60-80 % water in which macromolecules are contained [5]. The nucleus of a hydrogen atom consists of a single proton having positive charge. Also, a single electron with negative charge orbits the nucleus, making the whole atom electrically neutral.

An intrinsic property of nearly all elementary particles that proton, also, possesses is the **spin**. A proton, as a rotating mass, has angular momentum. Furthermore, taking

into account its electrical charge, proton has magnetic moment B and behaves like a small magnet. So, it is affected by external magnetic fields and when it moves induces a voltage in a receiver coil. Due to the size of the proton, we cannot look into hydrogen's nucleus thus we cannot see its intrinsic angular momentum or its spin. When we describe the rotation of a proton we are not referring to its angular momentum but to the visible motion of its magnetic axis B, which has the orientation of the proton's rotation axis.

When we expose our hydrogen nuclei to an external magnetic field, B_0, the nuclei undergoes the process of precession, where the external forces try to alter the orientation of its rotational axis, something commonly referred as wobble. This means, that our nuclei does not align with the field ,just, instantly, but undergoes precession, which happens at a characteristic frequency, which is proportional to the strength of the applied magnetic field and is called Larmor frequency (ω_0).

During this process the nuclei will gradually lose the energy stored as "wobbling" and eventually align with the applied magnetic field. The time needed for the nuclei to align depends on the decaying rate of the wobble phenomenon. Therefore, Larmor frequency is fundamentally important in our concept and of crucial significance to Magnetic resonance imaging.

The Larmor or precession frequency is the rate at which spins wobble when placed in a magnetic field [6].

$$\omega_0 = \gamma_0 B_0$$

where,

- ω_0 is the Larmor frequency in (MHz)
- γ_0 is the gyromagnetic ratio, is the ratio of a particle's magnetic moment to its angular momentum in (MHz/Tesla)
- B_0, is the strength of the magnetic field in (Tesla)
- and $\gamma_0 = 42.58$ (MHz/Tesla) which is the value for protons when exposed to a magnetic field at 1.5 T

But, what events take place in the prior process? Well, when the system has reached a stable state and the wobbling phenomenon has wore off, the longitudinal magnetization M_z vectors representing the individual magnetic moments in the z-direction are summarized. However, the nuclei , depending on their prior alignment, tend to align with the magnetic field parallel or anti-parallel. The direction of the alignment resides on which energy state is more favourable for the nuclei.

Under a steady-state condition, a slightly larger fraction aligns parallel to the Magnetic field. This slight difference is the one actually measured and is represented by the net magnetization vector. Due to the nature of this weak magnetization difference, the external magnetic field must be strong enough in order to obtain a detectable signal.

In order to induce energy to our stable spin system, we apply an electromagnetic wave (E/M) of the same frequency as the Larmor frequency. This is called resonance condition. We produce this E/M wave in a powerful radio transmitter and apply it on the object to be imaged by means of an antenna coil. This process is known as excitation of the spin system and during this process the axis is gradually tipped away from the z-axis to the x-y plane perpendicular to the direction of the main magnetic field.

The E/M wave is generated through a radio frequency (RF) pulse in order to be strong enough and is applied and sustained long enough, to tip the magnetization by exactly 90^o. This gives us a magnetization that is now denoted by M_{xy}. However, whenever a transverse magnetization rotates or precesses about the z-axis, it has the same effect of an electrical generator and induces an alternating voltage with the same frequency as the Larmor frequency: our MR signal. This particular signal is collected with high sensitivity sensors and processed via a computer to generate the MR Image. The whole procedure is graphically shown in (figure 2.2).

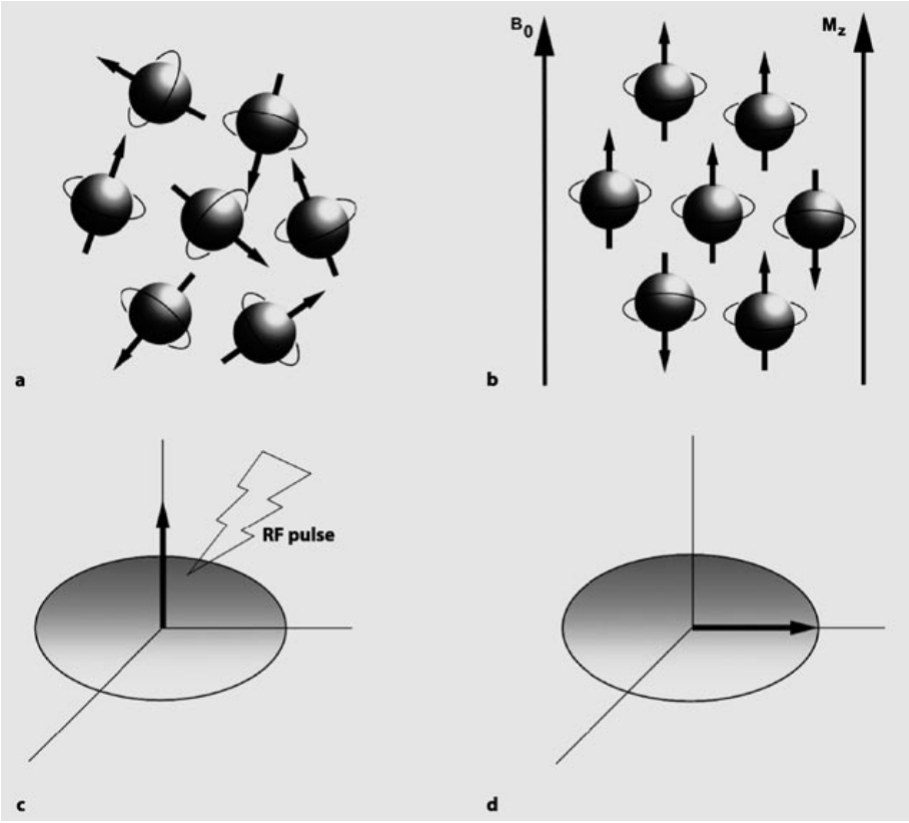


Figure 2.2: (**a**) Stable state of spin system, absence of magnetic field. (**b**) Production of longitudinal M_z magnetization on a magnetic field B_0. (**c**) Applying RF pulse. (**d**) From M_z longitudinal magnetization to transverse M_{xy}.

2.2 T1 and T2 Relaxation times

What is the process, though, to accumulate the MR signal?

Immediately after the excitation of the spins, the magnetization vector rotates to the

xy-plane and is now transverse magnetization (M_{xy}) just as we described previously. It is important to understand that the rotating transverse magnetization is what gives rise to the MR signal in the sensor coil. However, there are two independent phenomena that occur after excitation which affect the MR signal and force it to rapidly fade. The reason is, that they reduce the transverse magnetization and as a result they cause a return to the stable state present before excitation. These independent phenomena are called spin-lattice interaction and spin-spin interaction and they cause T1 relaxation and T2 relaxation, respectively.

- **T1: Longitudinal Relaxation**

As we mentioned above the transverse magnetization (M_{xy}) slowly decays and the magnetization vector and the magnetic moments gradually realign with the z-axis, which is the axis of the main magnetic field B_0. In this process the transverse magnetization slowly diminishes and while the projection of the magnetization vector onto the xy-plane (figure 2.3) decreases, the longitudinal magnetization, M_z, and its corresponding projection on the z axis, increase until the steady state conditions are restored [6]. Therefore, this process is known as longitudinal relaxation or T1 "recovery"

As it was stated before, during the process of excitation we have given an amount of energy to the nuclei to rotate its magnetization vector to the x-y plane, which means for the nuclei to undergo longitudinal relaxation it has to dissipate this excess energy to the environment. The time taken for T1 relaxation to occur vary and is dependant on the strength of the external magnetic field B_0 and the internal motion of the molecules. For example, biological tissues have T1 values between half to several seconds at 1.5 Tesla.

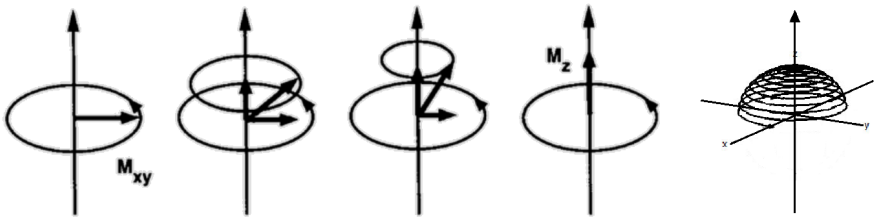


Figure 2.3: T1 Relaxation. Decay of M_{xy} and regrowth of M_z

- **T2/T2*: Transverse Relaxation**

To explain transverse relaxation, it is first necessary to explain the meaning of the "phase". Here, phase is represented as the angular difference of a magnetic moment's circular orbit. Think of two different and at this stage, independent nuclei A and B, that spin around the x-y plane with the same speed. If A began spinning before B did, then the A will be ahead of B for, let's say, 25 degrees. This means that the phase of A relative to B is +25 degrees and vice versa the phase of B relative to B is -25 degrees.

Immediately after the excitation process, part of the spins begin to spin simultaneously, meaning they have a phase of 0 degrees, or most commonly, they are in phase. This state is called phase coherence.

However, phase coherence is gradually lost, due to the fact that some spins speed up and some others slow down during their rotational movement, a phenomenon that we will soon explain. This, however, results in the decay of transverse magnetization M_{xy}, due to the fact that the individual magnetization vectors begin to cancel each other instead of adding together and the resulting sum becomes smaller and smaller until it diminishes and so does the MR signal (figure 2.4).

Simply put, transverse relaxation is the decay of transverse magnetization because the spins go out of phase. In contrast to longitudinal relaxation, transverse relaxation does not emit energy to the environment, instead the spins exchange this energy with each other.

Coherence is lost in two ways:

Energy transfer between spins as a result of local changes in the magnetic field. Such fluctuations are due to the fact that the spins are associated with small magnetic fields that randomly interact with each other. Spins precess faster or slower according to the magnetic field variations they experience. The overall result is a cumulative loss of phase. This is a purely spin to spin interaction and is bound to happen within a time constant T2 which is mostly independent of the strength of the external magnetic field B_0 or the focusing pulse we have emitted.

Time-independent inhomogeneities of the external magnetic field B_0. These are intrinsic inhomogeneities that are caused by the magnetic field generator itself and by the very person being imaged. They contribute to de-phasing, resulting in an overall signal decay that is even faster than described by T2. This phenomenon happens at the time constant of T2* which is always shorter that T2. The T2 effect is mostly a product of inhomogeneities at tissue borders , especially at air/tissue interfaces or are induced by local magnetic fields. The loss of the MR signal due to T2* effects is called free induction decay (FID). T2* effects can be concealed by using spin echo sequences.

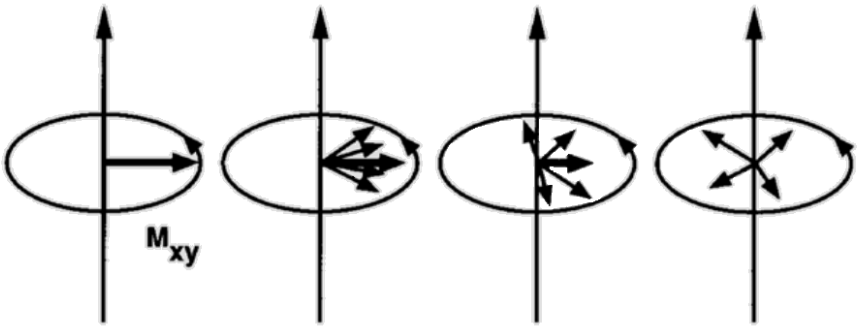


Figure 2.4: T2 Relaxation. Loss of phase coherence, loss of M_{xy}.

T2 denotes the process of energy transfer between spins, while T2* refers to the effects of additional field inhomogeneities contributing to de-phasing [6].

As we mentioned T1 and T2 relaxation processes are totally independent but they tend to occur simultaneously. It is worth noting that the MR signal has lost its power in the

first 100-300 ms due to T2 relaxation, which is long before T1 relaxation has completely restored the magnetization vector to the z axis.

2.3 Bloch Equations

Prior to water diffusion models it is reasonable to mention the mathematical form of the magnetic resonance phenomenon. The motion of the magnetization vector is described by Bloch in [7] using the equation.

$$\frac{d\mathbf{M}}{dt} = \gamma_0 \mathbf{M} \times \mathbf{B} - \frac{M_{xy}}{T_2} - \frac{M_z - M_0}{T_1} \tag{2.1}$$

where γ_0 is the gyromagnetic ratio, $\mathbf{B}$ the effective magnetic field, M_0 the equilibrium magnetization, T_1 and T_2 the relaxation times for the longitudinal and transverse magnetization, respectively.

We can write equation 2.1 after algebraic manipulations for the three spatial components as:

$$\dot{M}_x = \omega_0 M_y - M_x/T_2 \tag{2.2}$$
$$\dot{M}_y = -\omega_0 M_x - M_y/T_2 \tag{2.3}$$
$$\dot{M}_z = (M_0 - M_z)/T_1 \tag{2.4}$$

By representing the transverse magnetization equations (2.2) and (2.3), can be expressed in compact form as

$$M_{xy}(t) = M_x + iM_y \quad \text{and} \tag{2.5}$$
$$\dot{M}_{xy}(t) = M_{xy}(i\omega_0 - 1/T_2) \tag{2.6}$$

Thus, after integrating and imposing the initial condition $M_{xy}(0) = M_0$ we have:

$$M_{xy}(t) = M_0 e^{i\omega_0 t} e^{-t/T_2} \tag{2.7}$$

Similarly, after integrating the longitudinal magnetization equation 2.4 we have:

$$M_z(t) = M_0 + (M_z(0) - M_0) e^{-t/T_1} \tag{2.8}$$

The solutions of Bloch equation with respect to the direction of the magnetization vector are presented graphically in the (figure 2.5).

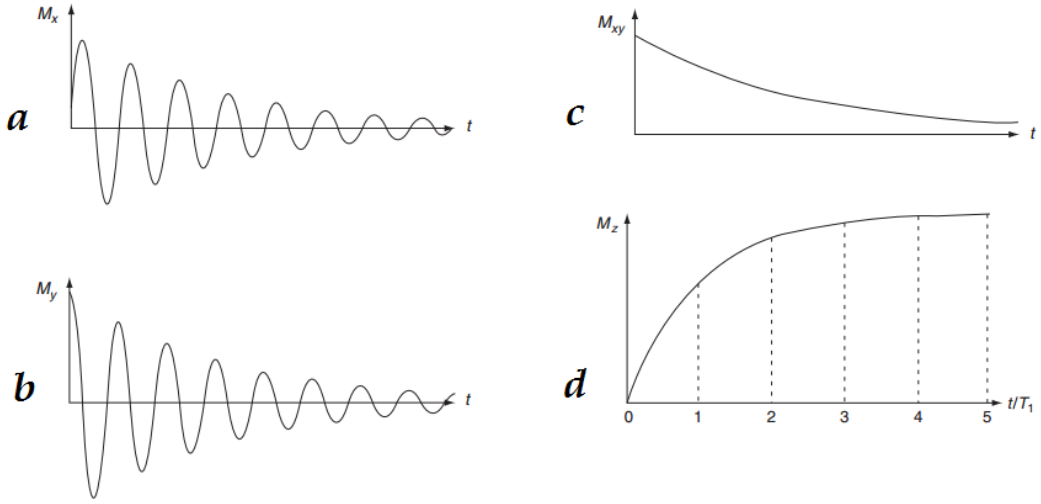


Figure 2.5: **(a)**,**(b)** Damped sine and cosine for the imaginary and real components (M_x, M_y) respectively.**(c)** Free induction decay signal M_{xy} with a time constant T2, **(d)** Exponential recovery of the longitudinal magnetization toward its equilibrium value with a time constant T1.

2.4 Image Contrast

Now, we will explain, briefly, what determines the contrast of an MR image and what we can do to manipulate it. There are three intrinsic features of a biological tissue that affect the image contrast. The proton density, T1 time and T2 time [6].

Proton density is the sum of the excitable spins per unit volume and it determines the maximum signal strength we can extract from the given tissue. We can choose the appropriate signal acquisition parameters to accentuate proton density contrast and minimize the effect of T1/T2 induced signal changes. By doing that we will take a proton-density weighted image, or simply proton density image.

T1 time, as we explained, is the time it takes for the excited spins to recover a certain percentage of their initial longitudinal magnetization. A safe 5×T1 time has passed can a tissue be available for a next excitation. Images with contrast that is mainly determined by T1 relaxation processes are called T1-weighted images (T1w).

T2 time generally represents how quickly or how much time we have before we receive the MR signal. It is the time we have before the MR signal decays. The T2 contrast of an MR image can be controlled and when the contrast is mainly influenced by spin-spin interaction then we have a T2-weighted image (T2w).

It is really important to mention that these intrinsic features vary widely from tissue to tissue. So does the image's contrast we can take from the very same tissue depending on which of these parameters we emphasize in the MR sequence. This provides the

basis for the exquisite soft-tissue discrimination and diagnostic potential of MR imaging. While, on a computed tomography (CT scan) this discrimination is virtually indistinct, MRI can distinguish them without the administration of a contrast medium.

Previously, it was mentioned that the main intrinsic features affecting the image contrast are the proton density and T1 and T2 times. Despite that, image contrast can be controlled by extrinsic contrast parameters, such as, **Repetition time (TR)** and **Echo time (TE)** that we will explain briefly in the next sections. Moreover, differences between water and fat will be analysed.

To examine the differences between fat and water, we have to proceed in molecular level. Molecular motion consists of rotational and transitional movements. Generally, when the molecular motion is fast, it is very difficult for a substance to release energy to its surroundings.

Fat consists of hydrogen atoms linked to carbon that make up large molecules. This fact results in a slow rate of molecular motion. Furthermore, fat molecules having low inherent energy, they are able to absorb energy efficiently.

In contrast, hydrogen atoms of water linked to oxygen, resulting in small and mobile molecules that have a high rate of molecular movement. Thus, water molecules have a high inherent energy and they are not able to absorb energy efficiently.

The differences between tissues that contain fat or water described above, are responsible for changes in image contrast. The reason why, arises from the different relaxation rates in each tissue.

- **Repetition Time (TR) and T1 Weighting**

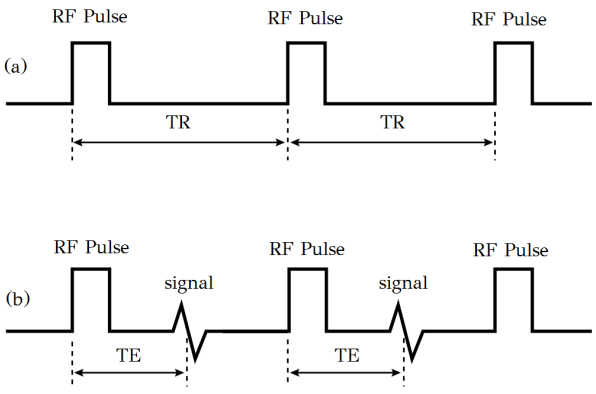


Figure 2.6: (**a**) Repetition time (TR). (**b**) Echo time (TE) [1].

For the formation of a successful MR image, tissue that is scanned must undergo several excitations and the resulting signals must be sampled many times. Repetition time (TR) is the time from the application of one RF pulse to the application of the next (figure 2.6(a)). TR is measured in (ms) and is crucial for the image contrast because it affects the length of a relaxation period after the application of one RF pulse to the beginning of the next [1].

Having a long TR results in more excited spins rotating back to the z-plane and contributing to the regrowth of longitudinal magnetization, which means that the more longitudinal magnetization can be excited by the next pulse, the stronger the MR signal will be emitted. On the other hand, choosing a short repetition time (approximately 600 ms), gives images where T1 effects are dominant. So, tissues with a short T1 relax quickly and hence give a strong signal after the next pulse and appear brighter on the image, while tissues with a long T1 receive only a little relaxation between two pulses and give less longitudinal magnetization on the next consecutive pulse, giving a weaker signal, and therefore appear darker. An example of two different tissue types (water and fat) is shown in (figure 2.7).

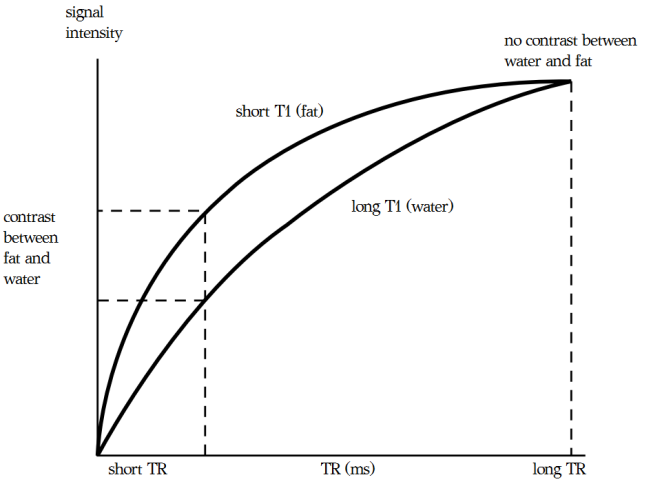


Figure 2.7: Differences of water and fat in T1 [1].

Figure 2.8: Sagittal T1 weighted image of spine. Intraspinal lipoma is bright as it contains fat. [1]

Choosing a fairly long repetition time, (which is approximately over 1500 ms) the majority of tissues, including those with long T1, return to equilibrium since they have enough time to relax and as a result they give similar signal strengths, which means that the effect of T1 on image contrast is minimal. With the above, we have shown that we can control the quality of T1 weighting of an MR image by choosing the repetition time.

The relationship between the MR signal of a tissue and its appearance on T1-weighted images is as follows: Tissues with a short T1 appear bright because they regain most of their longitudinal magnetization during the TR interval and thus produce a stronger MR signal (high signal intensity in T1-weighted images). Tissues with a long T1 appear dark because they do not regain much of their longitudinal magnetization during the TR interval and thus produce a weaker MR signal (low signal intensity in T1-weighted images). This dependence in TR is clearly shown in (figure 2.8) in which a lipoma in the spinal is depicted .

- **Echo Time (TE) and T2 Weighting**

The echo time dictates the degree of T2 relaxation time dependence on image contrast. Echo time, is the time between an Rf excitation pulse and the collection of the signal (figure 2.6 (b)). The TE affects the length of the relaxation period after the end of the RF excitation pulse and the amount of available MR signal to be collected by the receiver coil [1].

We have to point out here, that T2 relaxation time has the order of magnitude of hundreds milliseconds and is therefore far shorter than T1. The choice of TE is of high significance to the contrast of the image. If a short echo time is chosen, there are only small differences between different tissues, because T2 relaxation has only just began and there has been little difference in the time of the echo collection, giving us an image with a low T2 contrast.

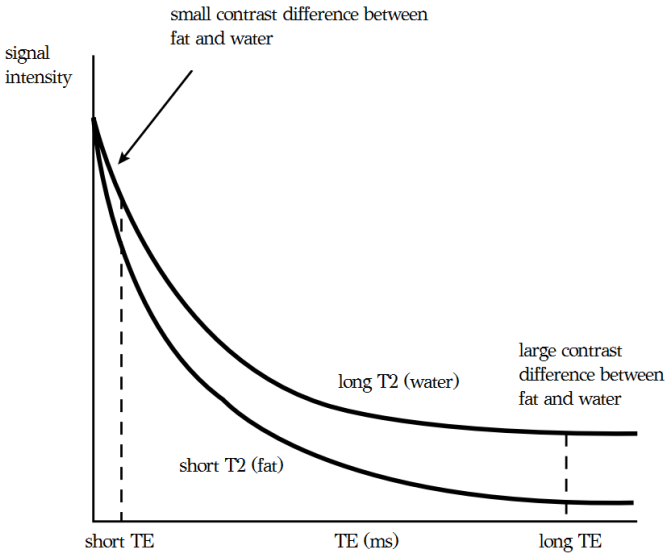


Figure 2.9: Differences of water and fat in T2 [1].

On the other hand, if a longer echo time is chosen, the resulting image accentuates differences between tissues of different T2 times. Tissues that have a short T2, lose most of their signal intensity and appear dark on the images while tissues that have a longer T2 haven't retained their magnetization and therefore appear brighter (figure 2.9).

Figure 2.10: CSF appears bright in a T2-weighted image

For instance, cerebrospinal fluid (CSF) is brighter compared with brain tissue , in T2 weighted images, due to the fact that CSF is mostly water and water has a longer T2. Thus, with the selection of an appropriate echo time TE, we can manipulate the degree of T2 weighting of the MR image (figure 2.10).

As a conclusion, tissues with a short T2 appear dark on T2-weighted images, while, tissues with a long T2 appear bright on T2-weighted images!

2.5 The Diffusion of water

The main target of diffusion weighted imaging (DWI) is to give insight into the diffusion of water molecules, a complex process in many biological and physical systems. DWI has significant clinical applications. Specifically, it offers high sensitivity of detecting acute ischaemic damage in the brain and supports clinical faculty in stroke management. In this chapter basic diffusion concepts such as magnetic resonance measurement of diffusion and anisotropic diffusion will be discussed in order to present the IVIM model.

2.5.1 Isotropic Diffusion - Basic Diffusion Concepts

The mobility of the molecules (molecular diffusion or Brownian motion) in their micro environment in a fluid, is reflected by a characteristic self-diffusion constant **D** [8]. This motion was described by Einstein in 1905 for pollen on water surface. From one hand (Fick 1855) for the quantitatively description of the diffusion process, asserted that the rate of flow of the diffusive substance in a particular direction would be proportional to the concentration gradient in that direction. Fick's first law of diffusion states that the diffusive process spontaneously drives the diffusive substance from areas of higher concentration to areas of lower concentration and expressed as:

$$F = -D\frac{\partial C}{\partial x} \tag{2.9}$$

where **F** is the rate of transfer of the diffusing substance through the unit area of each section of the sample, **C** is the concentration of the diffusing substance, **x** is the space coordinate and **D** is the diffusion coefficient.

Fick's second law of diffusion is derived from 2.9 using a conservation of tracer equation $\partial C/\partial t = -\partial F/\partial x$ which states that the local rate of increase in tracer concentration equals to the rate at which the flow decreases with distance. Thus, Fick's second law describes the diffusion process in one dimension in terms of the temporal and spatial partial derivatives of C:

$$\frac{\partial C}{\partial t} = D\Big(\frac{\partial^2 C}{\partial x^2}\Big) \tag{2.10}$$

By generalizing 2.10 in three spatial coordinates the global behaviour of isotropic diffusion can be expressed as:

$$\frac{\partial C}{\partial t} = D\Big(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} + \frac{\partial^2 C}{\partial z^2}\Big) = D\nabla^2 C \tag{2.11}$$

The term "isotropic" premises unrestricted molecule motion with no preferential direction for the diffusive motion.

Instead of describing self-diffusion (random walk) process in a homogeneous medium in terms of the concentration C, we introduce a function that describes the probability P of finding a particle in position x at time t. For instance, we want to follow the displacements of a molecule positioned at $X = A$ at time $t_0 = 0$ and want to compute the probability of finding it at $X = B$ at time t. Thus we want to find $P(X = x_B|X = x_A, t)$ but in our case, the self-diffusion, the probability is independent of the starting point. Consequently, we define a vector for the relative dynamic displacement, $X = (x_B - x_A)$ to formally describe the process with $P(X, t)$. According to Fick's second law, the self-diffusion mechanism can be expressed as:

$$\frac{\partial P}{\partial t} = D\nabla^2 P. \tag{2.12}$$

In a time scale of few milliseconds, water molecules are displaced between 10^{-8} and 10^{-4} m when exposed in a magnetic field. To visualize the displacement of molecules in a voxel we define the displacement distribution which describes the proportion of molecules that undergo displacement in a specific direction and distance. Most water molecules are assumed to travel short distances within a voxel, while, a few of them travel farther. This fact lead us to assume a Gaussian distribution.

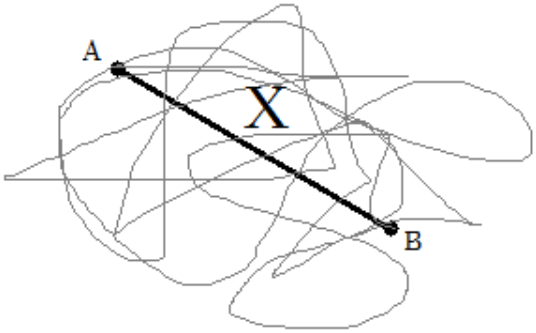


Figure 2.11: Random Walk of a particle

$$P(X, t) = \frac{1}{\sqrt{4\pi Dt}} \quad e^{\frac{-X^2}{4Dt}}. \tag{2.13}$$

The spread of the Gaussian is characterized by the variance ($\sigma^2 = 2Dt$) and the mean square dynamic displacement for the net vector of distance $\langle X^2 \rangle$ travelled by a molecule in time t is:

$$\langle X^2 \rangle = \int_{-\infty}^{+\infty} X^2 P(X, t) dX = 6Dt \tag{2.14}$$

At 37^{o}C (which is the normal body temperature) the diffusion coefficient of water is approximately $D = 3{\cdot}10^{-9} mm^2/sec$ [9]. From the other hand one can model self-diffusion as random walk (fig 2.11) starting from point A to B with X the net displacement vector .

2.5.2 MR Measurement of diffusion

This section is focused on how can MRI signal be sensitized to water diffusion. In the previous chapter we described that, MRI image is created by magnetic properties of protons. Supposing that we have a spin, that travels along an arbitrary path $r(t)$, through a gradient field by the waveform, $G(t)$, this leads to phase accumulation, ϕ [2].

$$\phi(t) = -\gamma_0 \int_0^t \overrightarrow{G}(t') \cdot \overrightarrow{r}(t')dt' \tag{2.15}$$

In a medium, with a large number of molecules and consequently, spins, it is logical to foresee there are many different magnetic moments; either positive or negative. Thus, the net phase shift tends to zero. Generally, the directional motion of spins through a gradient field results in a measurable phase. In contrast, a great percentage of randomly travelling spins provides no net phase shift, but they contribute to dephasing as signal attenuation. In addition, the degree of signal attenuation increases along with the increment of the variance of the phase shifts distribution. So, from 2.15 we conclude to phase shift increment, as the time interval, the gradient strength and arbitrary traveling path are increasing. This phenomenon was formally described and proved by Torrey in 1956 [10] to modify the Bloch equation with diffusion terms by adding the term $\nabla D \nabla (M_z - M_0)$. The solution of the Bloch-Torrey equation then is:

$$S(t) = S_0 e^{-D \int_0^t \overrightarrow{k}(t') \cdot \overrightarrow{k}(t')dt'} \tag{2.16}$$

where: $S(t)$ is the diffusion signal from the MRI scanner, S_0 is the signal without diffusion gradient, D is the diffusion coefficient and $k(t)$ relates to the time integral of the gradient waveform which is:

$$\overrightarrow{k}(t) = \gamma_0 \int_0^t \overrightarrow{G}(t')dt'. \tag{2.17}$$

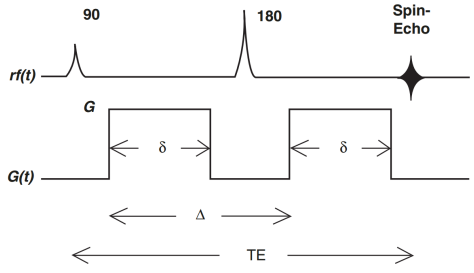


Figure 2.12: Standard pulsed field gradient waveform.

Equations 2.16 and 2.17 are written for the general case of an arbitrary gradient waveform G. Usually the most used sequence in practice, is the Stejskal-Tanner pulsed field gradient sequence (PFG), (figure 2.12). The integral in 2.16 is refereed in bibliography as the b-value or b-factor as introduced by Le Bihan in [11]. By applying the waveform in (figure 2.12) to (2.17) and the integral in (2.16), we have:

$$b = \int_0^{TE} \overrightarrow{k}(t') \cdot \overrightarrow{k}(t')dt' = (\gamma_0 G\delta)^2[\Delta - \frac{\delta}{3}] \tag{2.18}$$

In practice, altering the amplitude or duration of the gradient (and, thus, the b-value) in a diffusion imaging sequence, results in a change of signal as a function of the b-value, under the rule

$$S(b) = S_0 e^{-bD}. \tag{2.19}$$

In figure (2.13), the signal (MR image) as a function of b-value is shown. It is clear that, when b-value increases, leads to signal loss, unrestricted molecules (CSF) loose MR signal faster than molecules in restrictive environment (brain parenchyma). As mentioned previously this is result of spin dephasing.

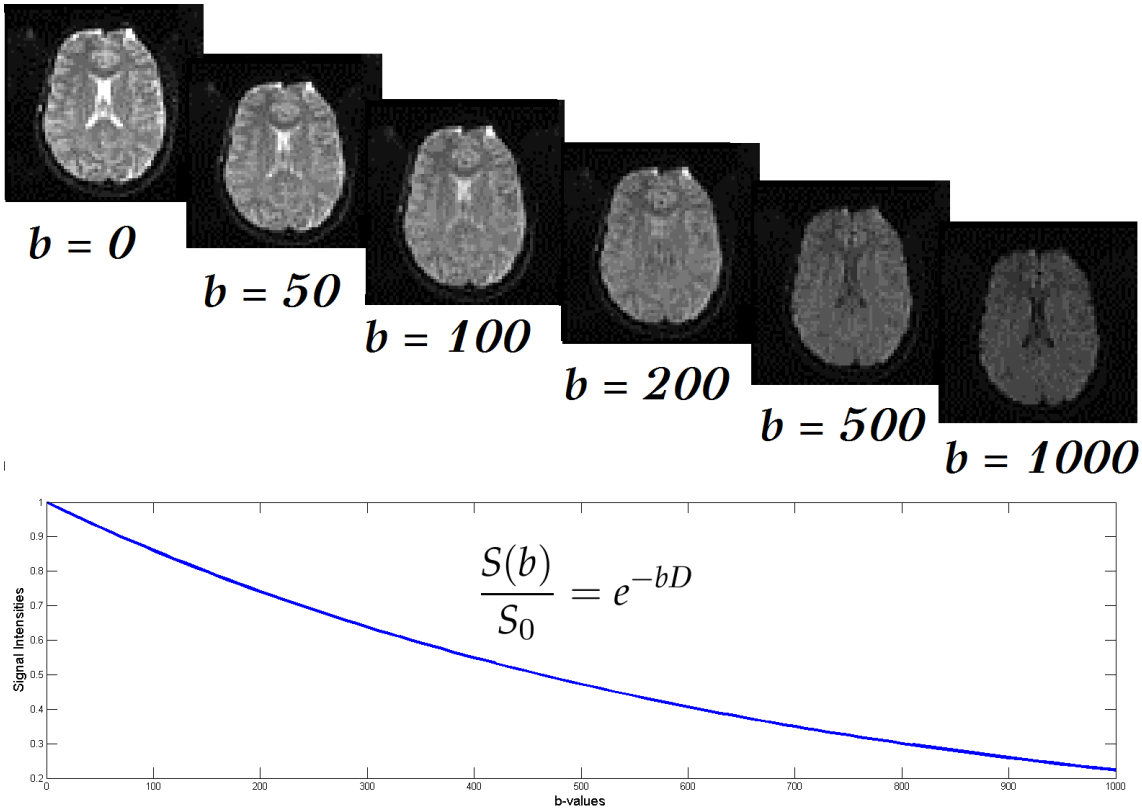


Figure 2.13: Signal as a function of the b-value

In order to calculate the diffusion coefficient **D** from (2.19), we need at least a diffusion sequence with two different b-values. This procedure is repeated for every pixel on a Diffusion MR image. When there is no diffusion gradient ($b = 0$), the signal is indicated as $S(b) = S_0$, meaning no signal loss and no diffusion sensitivity of the acquisition. So, in the simplest case of determining **D** with 2 b-values (b_1, b_2), we have:

$$S(b_1) = S_0 e^{-b_1 D} \tag{2.20}$$

and

$$S(b_2) = S_0 e^{-b_2 D}. \tag{2.21}$$

By taking the ratio of both sides of the previous equations and after that, the logarithm of both sides we conclude that:

$$D = \frac{1}{(b_2 - b_1)} ln \left[\frac{S(b_1)}{S(b_2)} \right]. \tag{2.22}$$

Due to the fact that D is dependent on the direction of diffusion and the b-value, the physical diffusion coefficient is replaced by the apparent diffusion coefficient (**ADC**). On the other hand, in clinical practice, more than two b-values are acquired for assessing in a more accurate way the exponentially decay of the diffusion signal. This fact leads to non-linear fitting to calculate the diffusion coefficient. In section 3.1, Non-Linear Least Squares are presented for the calculation of **ADC**.

2.5.3 Anisotropic Diffusion

In the previous section, we modelled the process of diffusion of molecules as a Gaussian (isotropic) distribution, meaning that spins have no preference in direction as they travel through a medium or tissue in one dimension. Mobility of water molecules in brain white matter is mostly anisotropic as water preferentially moves along white matter fibers than across them as myelin inhibits free diffusion [12].

Diffusion Tensor Imaging is a novel technique based on diffusion theory that can represent 3D anatomical structures in the human brain, and specifically white matter tracts [8]. It has been used for pre-operative brain MRI in order to assert integrity of tracts in the vicinity of the tumour size as well as for other brain pathologies such as, multiple sclerosis, dyslexia and schizophrenia.

The diffusion tensor (D) consists of nine elements (see 2.23). The diagonal elements D_{xx}, D_{yy}, and D_{zz} define the diffusion coefficients along the main system of axons xyz of the MRI scanner. Non-diagonal terms of (2.23) represent the offset of principal direction of diffusivity from the reference frame (fig 2.14). In case the reference frame coincides with the self-direction of diffusion, non-diagonal terms are equal to zero. Furthermore, for uncharged molecules, tensor D is symmetric. Thus, only six elements are required.

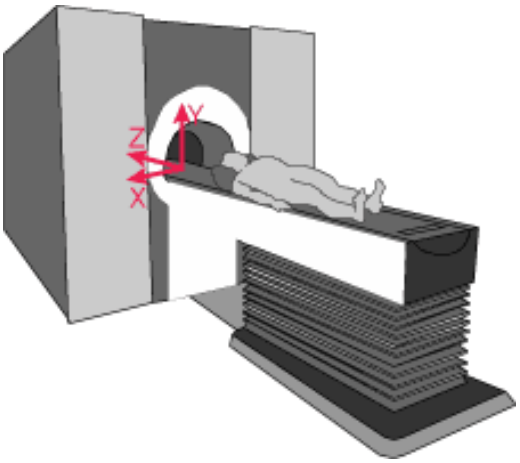


Figure 2.14: MR scanner x, y, z co-ordinate system. Figure taken from: Gardner Lab. Neurosciences Institute, Stanford University

$$D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \tag{2.23}$$

2.5.4 Calculation of Diffusion Tensor

The tensor basis solution to Bloch equations with diffusion terms, with respect to the Stejskal-Tanner pulsed field gradient sequence for (DTI) is [13]:

$$S_i = S_0\, e^{-b\, g_i^T\, D\, g_i} \tag{2.24}$$

where: D is the diffusion tensor and $g_i = [G_{xi}\, G_{yi}\, G_{zi}]^T, \quad i = 1, ..., N$ are the gradient direction pulses. For the directional sampling, usually $N = 6$, $N = 12$, and $N = 42$ uniformly distributed directions of measurements, are used. In order to succeed better accuracy on the direction of molecules, we set $N = 162$ or $N = 642$ but it is time consuming especially for the object to be imaged. So, in order to find D, from 2.24 we conclude that, the minimum images required are seven (six images along different directions and one with no diffusion gradient).

After a series of algebraic manipulations on 2.24 we have:

$$-\frac{ln\left(\frac{S_i}{S_0}\right)}{b} = G_{xi}^2 D_{xx} + G_{yi}^2 D_{yy} + G_{zi}^2 D_{zz} + 2G_{xi}G_{yi}D_{xy} + 2G_{xi}G_{zi}D_{xz} + 2G_{yi}G_{zi}D_{yz} \tag{2.25}$$

In the process of finding the elements of the tensor, we write D as a column vector $\mathbf{d} = [D_{xx}\, D_{yy}\, D_{zz}\, D_{xy}\, D_{xz}\, D_{yz}]^T$, we introduce the **H** matrix and the right hand side **F**.

$$\mathbf{H} = \begin{bmatrix} G_{x1}^2 & G_{y1}^2 & G_{z1}^2 & 2G_{x1}G_{y1} & 2G_{x1}G_{z1} & 2G_{y1}G_{z1} \\ G_{x2}^2 & G_{y2}^2 & G_{z2}^2 & 2G_{x2}G_{y2} & 2G_{x2}G_{z2} & 2G_{y2}G_{z2} \\ G_{x3}^2 & G_{y3}^2 & G_{z3}^2 & 2G_{x3}G_{y3} & 2G_{x3}G_{z3} & 2G_{y3}G_{z3} \\ & & & \vdots & & \\ G_{xN}^2 & G_{yN}^2 & G_{zN}^2 & 2G_{xN}G_{yN} & 2G_{xN}G_{zN} & 2G_{yN}G_{zN} \end{bmatrix} \tag{2.26}$$

$$\mathbf{F} = -\left[\frac{ln\left(\frac{S_1}{S_0}\right)}{b}, \frac{ln\left(\frac{S_2}{S_0}\right)}{b}, \ldots, \frac{ln\left(\frac{S_N}{S_0}\right)}{b}\right]^T \tag{2.27}$$

Hence, the diffusion tensor is computed by solving the linear system,

$$\mathbf{Hd} = \mathbf{F}. \tag{2.28}$$

In order to present the Diffusion Tensor, a healthy patient's axial slice of the brain with $N = 12$ uniformly distributed gradient directions has been used (figure 2.15). The images below, were obtained from an 1.5 Tesla Siemens MR scanner of University General Hospital of Heraklion "PAGNI".

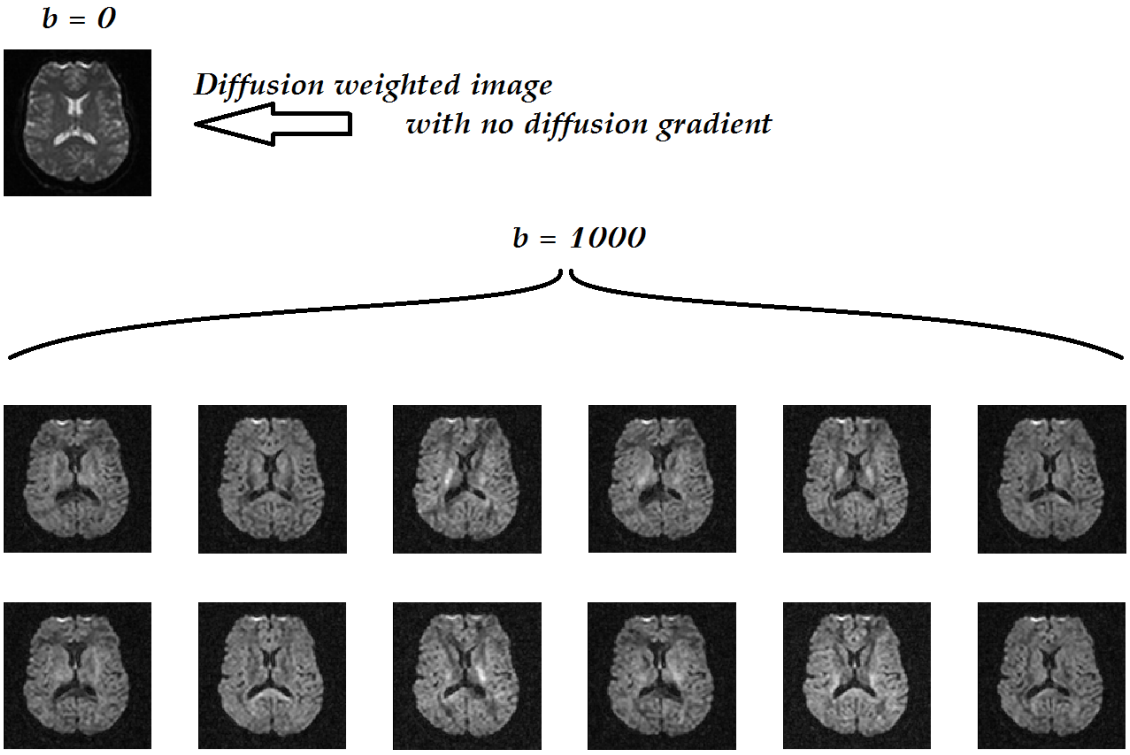


Figure 2.15: DTI Data: 13 diffusion weighted images: one with no diffusion gradient (b = 0) and 12 images with uniformly distributed gradient directions with (b = 1000).

Hence, by solving the linear system 2.28 for every pixel on the image, the Diffusion Tensor is presented in figure 2.16 . Every image, represents the corresponding element of the tensor.

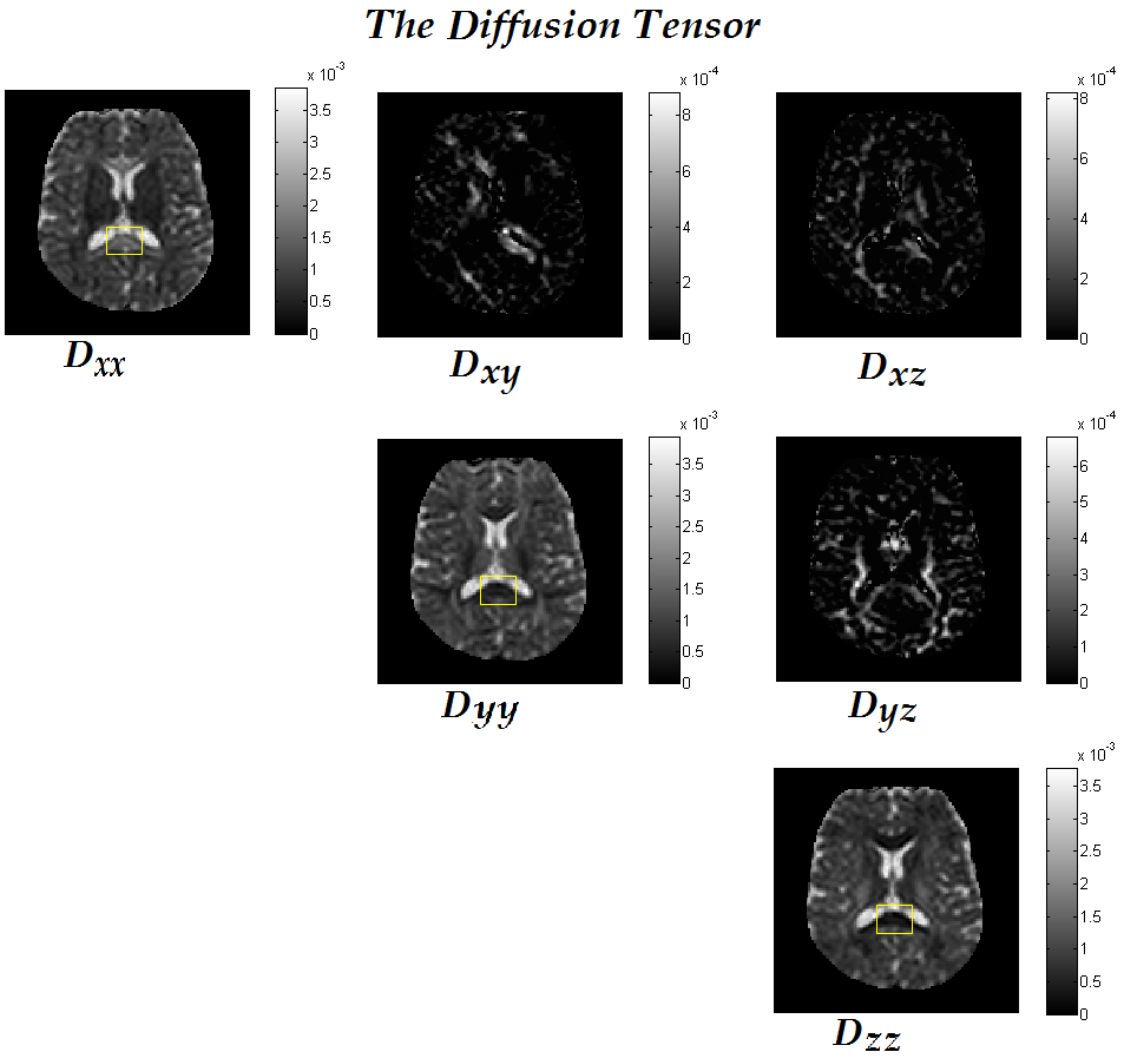


Figure 2.16: Calculation of the diffusion tensor. On the diagonal images that represent the diffusion through the xx, yy, and zz axons respectively, high diffusion coefficient in xx (yellow box) is observed, which is in contrast to the case of diffusivity along yy and zz. That happens due to the alignment of callosal fibers with the direction of diffusivity along xx.

2.5.5 Visualization of the Diffusion Tensor

After calculating the diffusion tensor, it is possible to visualize the direction of water diffusion through a medium (brain tissue). The eigenvalues (λ_1, λ_2, λ_3) and eigenvectors (e_1, e_2, e_3) of the tensor D are of great interest, which describe the apparent diffusivities and the directions along the axes of principle diffusion [14] .

The diffusion tensor can be visualized using an ellipsoid with the eigenvectors defining the directions of the principle axes and the ellipsoidal radii defined by the eigenvalues (figure 2.17). In the present work, considering the eigenvalues in descending order, λ_1 represents the direction with the highest diffusion coefficient.

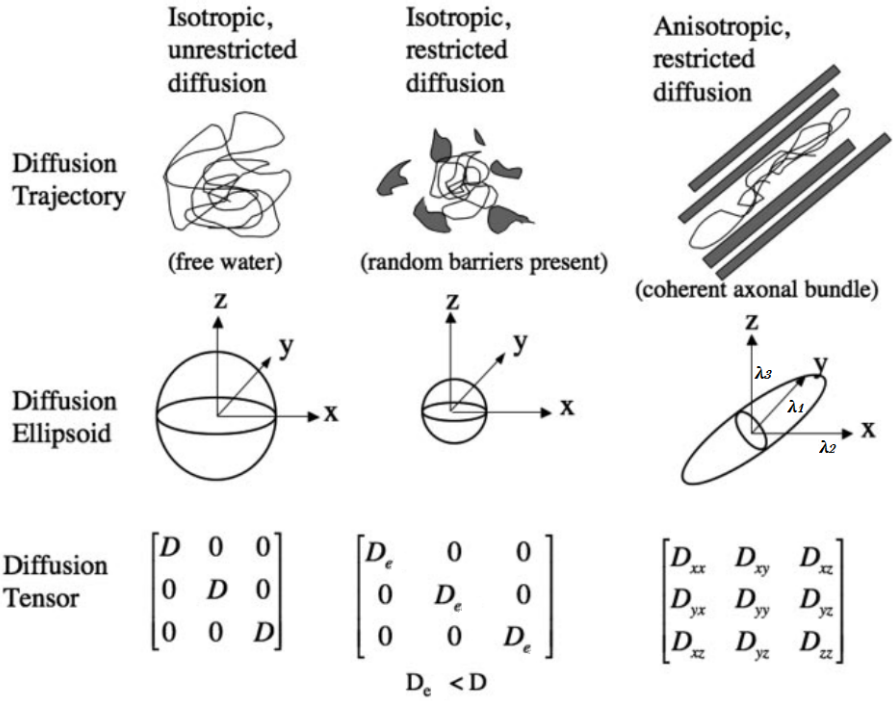


Figure 2.17: Diffusion ellipsoids. Figure by [15]

The diffusion ellipsoids are of the form:

$$\frac{x^2}{2\lambda_1} + \frac{y^2}{2\lambda_2} + \frac{z^2}{2\lambda_3} = 1 \tag{2.29}$$

The visualization of the diffusion tensor with the diffusion ellipsoids is shown in (figure 2.18).

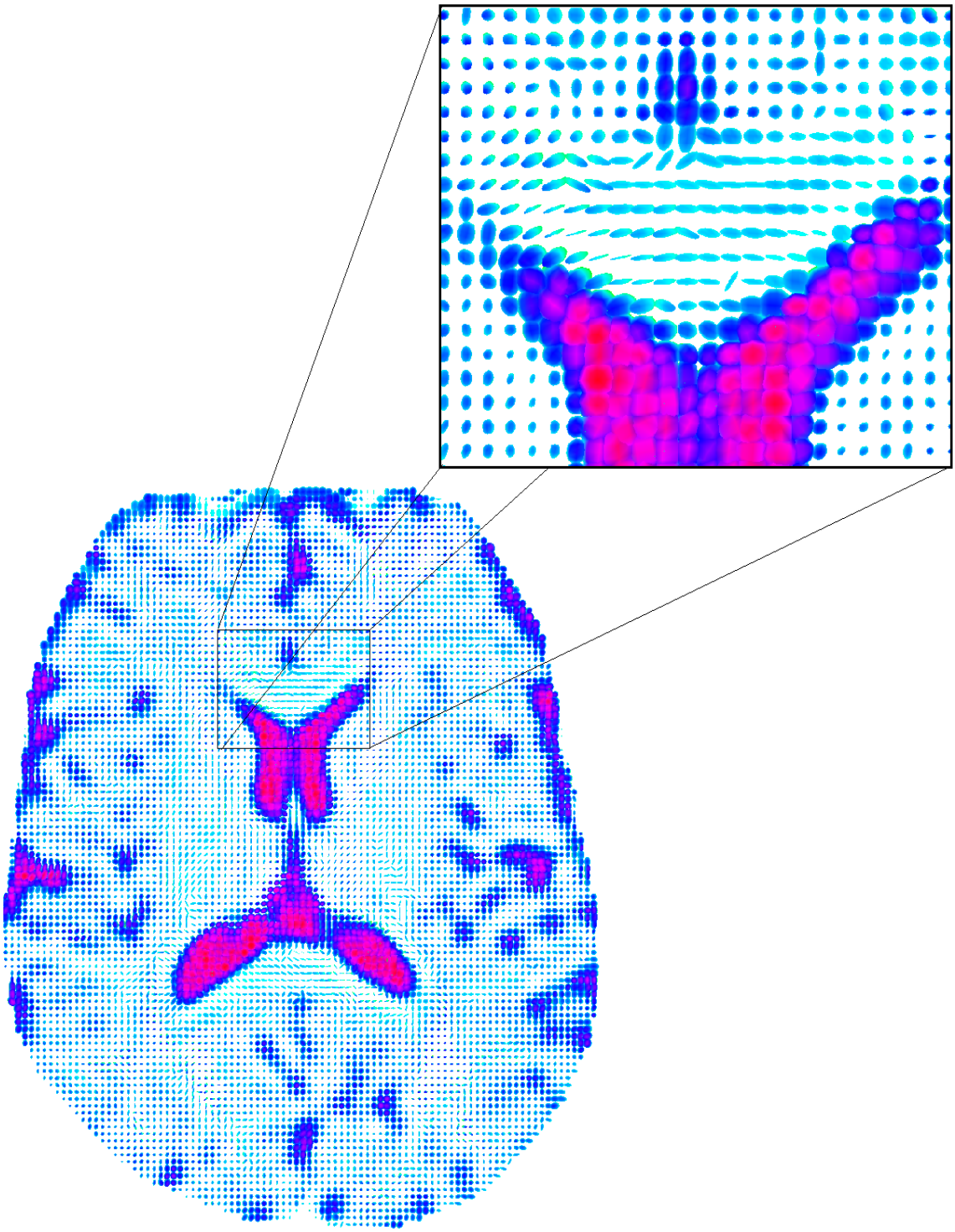


Figure 2.18: Diffusion ellipsoids map. Black box: both anisotropy in diffusion in the genu of corpus callosum and isotropy in CSF are clearly shown.

Since the interpretation of the DT is a hard and time consuming procedure, a simplification of the data would be necessary. Hence, it is reasonable to search for other measures

(bio-markers) that can produce information to radiologists or doctors. Other measures that derived from DTI are: Fractional Anisotropy (FA), Mean Diffusivity (MD), Radial Diffusivity (RD) and Volume Ratio (VR).

Firstly, for an overview of diffusion in a voxel, the most commonly used bio-marker, Mean Diffusivity, is proposed. MD characterizes the mean-squared movements of molecules and diffusion restrictions in tissue (2.30). Radial diffusivity is calculated by averaging the orthogonal components of the principal eigenvector. RD is associated with loss of myelination and cross-sectional axonal injuries (2.31). Also it witnesses changes related to the aging process.

Fractional Anisotropy and Volume ratio are two scalars that define the diffusion anisotropy. FA measures the fraction of diffusion that is anisotropic by looking at the normalized variance of the eigenvalues in the tensor (2.32). $FA = 0$ indicates isotropic diffusion while $FA = 1$ fully anisotropic. VR stands for the ratio of the ellipsoid volume to the volume of a sphere of radius $\langle\lambda\rangle$ (2.33). As FA, $VR's$ range is $[0, 1]$ but, $VR = 0$ stands for anisotropy while $VR = 1$ isotropy.

$$MD = \langle\lambda\rangle = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \tag{2.30}$$

$$RD = \frac{\lambda_2 + \lambda_3}{2} \tag{2.31}$$

$$FA = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{2.32}$$

$$VR = \frac{\lambda_1\lambda_2\lambda_3}{\langle\lambda\rangle^3} \tag{2.33}$$

In (figure 2.19), scalar measures of diffusion anisotropy are presented (FA, VR) as well as measures (MD, RD) that give information about mean diffusivity and changes to myelination.

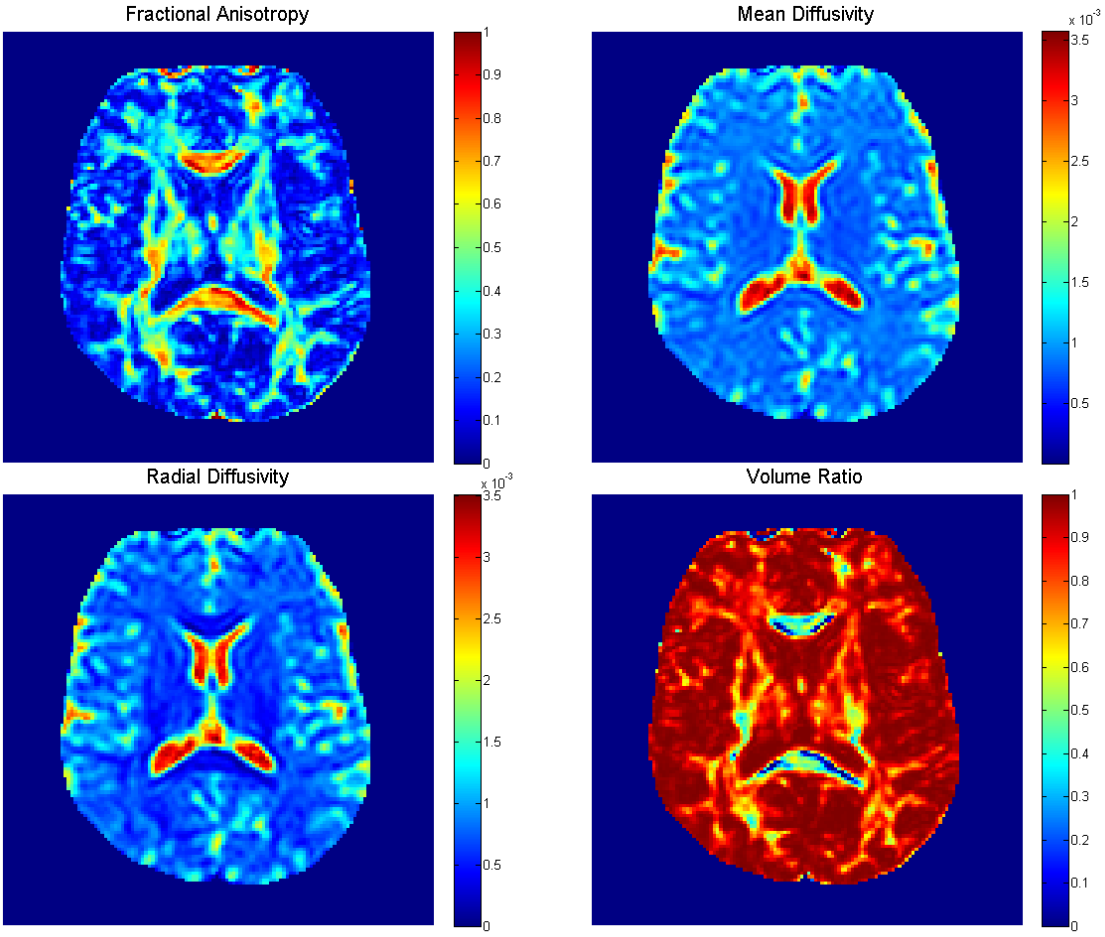


Figure 2.19: Fractional Anisotropy, Mean Diffusivity, Radial Diffusivity, and Volume Ratio maps

2.5.6 Intravoxel Incoherent Motion (IVIM)

The above DTI scalar measures, have been used to characterize abnormalities and diseases in the human brain. To examine water diffusivity in other anatomical areas, the simple ADC model (2.19) is proposed. From this mono-exponential decay, the apparent diffusion coefficient (ADC) could be calculated quantitatively by linear fitting to the logarithmic scale of the signal intensities, using only two b-values.

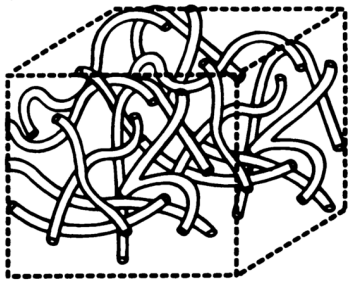


Figure 2.20: Capillary network in a voxel.

Practically, by the use of several b-values, many studies have experimentally concluded that DWI signal cannot be well described by the mono-exponential model (2.19), especially on highly vascularized organs like the liver and the pancreas [16]. The DWI signal on these well perfused organs, is also influenced by the tissue micro-structure, hence the microcirculation of blood in the capillary network (figure 2.21).

Intravoxel Incoherent Motion (IVIM) imaging method was developed for the visualization of microscopic translations of water. The main idea in [17] was that the signal in a diffusion MR image could be contaminated by the microcirculation of blood (perfusion) in the capillary network in a voxel. This method is capable of indicating the contribution of perfusion to the diffusion coefficient D and permits separate evaluation of parametric maps related to the aforementioned phenomena.

Due to the pseudo-random organization and the high quantity of capillaries (5700 per cubic millimeter) of the capillary network in a voxel, microcirculation of blood was considered as an incoherent motion (figure 2.20). The model that proposed for IVIM is:

$$\frac{S_b}{S_0} = (1-f)e^{-bD} + fe^{-bD^*} \tag{2.34}$$

where: f is the micro-perfusion factor, $f \in [0,1]$. It denotes the ratio of water flowing in capillaries to the total water contained in the voxel. D is the diffusion coefficient and D^* is the pseudo-diffusion coefficient.

Expected values of D^* coefficient are approximately one order of magnitude greater than D [18]. Further analysis, of this model is presented in the next chapter.

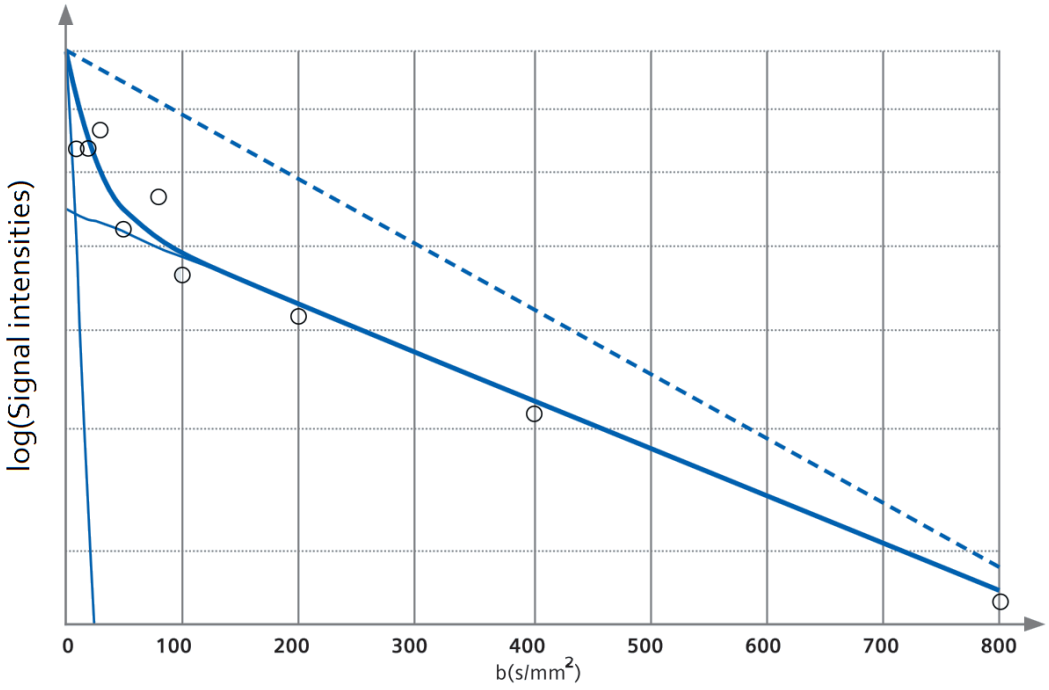


Figure 2.21: A signal decay in semi-logarithmic scale as a function of the b-values. Circles correspond to the semi-logarithmic transformed DWI signal intensities. The bold solid line is the IVIM fitted curve providing D, D* and f. Light solid lines represent D and D* decay curves respectively. The dashed line is the mono-exponential fit providing ADC [16]

3

Diffusion Weighted Imaging - IVIM Analysis

This section is focused on the materials and methods of the presented DWI analysis as well as, on the technical aspects required for providing the diffusion parameters from the IVIM model.

3.1 Mathematical Framework of IVIM Computation

The aforementioned DWI parameters (f, D, D^*) from the IVIM model in the previous chapter, are estimated by the non-linear inverse problems theory. One of the most preferred techniques of this theory is, the non-linear least squares. Suppose we have data points $(b_1, y_1), (b_2, y_2), ..., (b_n, y_n)$ and a given model function $G(\mathbf{x}, b)$ (figure 3.1). For the parameters $x = [x_1, x_2, ..., x_m]$ we want to find a local minimizer **x*** for

$$F(x) = \frac{1}{2}\sum_{i=1}^{n} r_i^2(x) = \frac{1}{2} \parallel r(x) \parallel^2 = \frac{1}{2} r(x)^T r(x), \tag{3.1}$$

known as objective function. Firstly we assume that there exists a $x\prime$ such that $y_i = G(x\prime, b) + \epsilon_i$, where, $\{\epsilon_i\}$ are the errors of the measurements y_i and for any x the residuals $r_i : \mathbb{R}^m \to \mathbb{R}$ with $r_i(x) = y_i - G(x, b) \quad i = 1, ...n \quad n \geq m$.

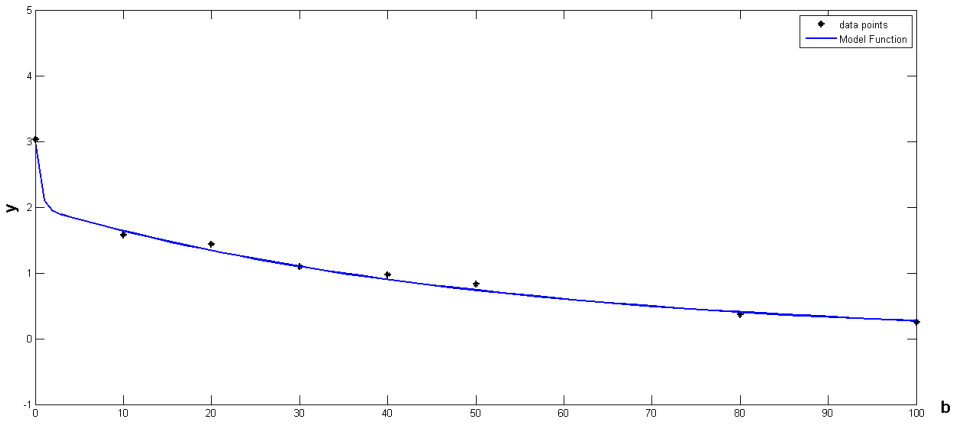


Figure 3.1: Data points marked by (*) and model $G(\mathbf{x}, b)$ by full line.

- **Definitions and Notation**

In order to comprehend the methods of finding a minimizer for (3.1), some useful definitions and theorems are presented. In general, the main problem is to find a global minimizer as stated below.

Definition 1 (Global minimizer) *Given* F: $\mathbb{R}^m \to \mathbb{R}$, *find* $\quad x^+ = min_{x \in \mathbb{R}^m} F(x)$.

Finding a global minimizer is a very difficult procedure hence, a simplification of the problem is to search for a local minimizer (x^*) in a neighbourhood of x with radius δ in definition 2.

Definition 2 (Local minimizer) *Given F:* $\mathbb{R}^m \to \mathbb{R}$, *find* x^* *such that* $F(x^*) < F(x)$ *for* $\| x - x^* \|_2 < \delta, \quad \delta > 0$.

Assuming that the objective function F is differentiable and smooth, the truncated Taylor expansion of $F(x) = L_n(h) + O(h^{n+1}), \quad n = 1, 2, ...$ around x is:

$$F(x+h) = L_2(h) = F(x) + h^T g + \frac{1}{2} h^T H h + O(\| h \|^3) \tag{3.2}$$

where g is the gradient and H is the Hessian,

$$g = F'(x) = \begin{bmatrix} \frac{\partial F}{\partial x_1}(x) \\ \vdots \\ \frac{\partial F}{\partial x_m}(x) \end{bmatrix}, \qquad H = F''(x) = \left[\frac{\partial^2 F}{\partial x_i \partial x_j}(x) \right]. \tag{3.3}$$

For simplification, we need formulas for the derivatives of F. Supposing that r has continuous second partial derivatives, the truncated Taylor expansion of r is:

$$r(x+h) = l_1(h) = r(x) + J(x)h + O(\| h \|^2), \tag{3.4}$$

with, $\mathbf{J} \in \mathbb{R}^{n \times m}$ to be the Jacobian matrix.

$$(J(x))_{i,j} = \frac{\partial r_i}{\partial x_j}(x), \quad i = 1, ..., n, \quad j = 1, ..., m. \tag{3.5}$$

Moreover, from 3.1 the gradient of $F : \mathbb{R}^m \to \mathbb{R}$ is:

$$F'(x) = \frac{\partial F}{\partial x_j}(x) = \sum_{i=1}^{n} r_i(x) \frac{\partial r_i}{\partial x_j}(x) = J(x)^T r(x). \tag{3.6}$$

Finally, from 3.6, the element of the Hessian of F at position (j, k) is:

$$\frac{\partial^2 F}{\partial x_j \partial x_k}(x) = \sum_{i=1}^{n} \left(\frac{\partial r_i}{\partial x_j}(x) \frac{\partial r_i}{\partial x_k}(x) + r_i(x) \frac{\partial^2 r_i}{\partial x_j \partial x_k}(x) \right). \tag{3.7}$$

In matrix form, one can see that, the Hessian is:

$$F''(x) = J(x)^T J(x) + \sum_{i=1}^{n} r_i(x) r''(x). \tag{3.8}$$

Usually, when the residuals are significantly smaller than the first term of the Hessian,

the Gauss-Newton approximation of the Hessian (H_{GN}) is used:

$$F''(x) = H_{GN} = J(x)^T J(x). \tag{3.9}$$

A necessary condition for a local minimizer, x^*, is :

$$g^* = F'(x^*) = 0. \tag{3.10}$$

A point (x_s) that satisfies the necessary condition is called stationary. So, a stationary point may be a local minimizer or a local maximizer. The next theorem ensures that x_s is a local minimizer.

Theorem 1 (Sufficient condition for a local minimizer) *Assume that x_s is a stationary point and $F''(x_s)$ is positive definite. Then, x_s is a local minimizer.*

Non linear optimization methods are iterative. Given a starting point (x_0) methods produce a sequence of vectors $x_1, x_2, ...$ which converges to x^* (the local minimizer). The choice of the starting point (x_0) is critical and may affect the result, specially when the objective function has many minimizers.

To be sure that the methods are going to converge in a local minimum (x^*) and not in a local maximum, the descending condition must be satisfied:

$$F(x_{k+1}) < F(x_k). \tag{3.11}$$

Also, the error sequence $\{e_k\}$ that is produced by the iterative method, must be decreasing, eventually, i.e.

$$\| e_{k+1} \| < \| e_k \|, \quad k > K \in \mathbb{N}. \tag{3.12}$$

In every iteration, the next direction (**h**) that leads us to the closest minimizer x^* must be calculated. By taking the Taylor around x and with direction h, we have:

$$F(x + \alpha h) = F(x) + \alpha h^T F'(x) + O(\alpha^2) \tag{3.13}$$
$$\simeq F(x) + \alpha h^T F'(x) \tag{3.14}$$

If $F(x + \alpha h)$ is a decreasing function of α, then, h said to be the descent direction and concludes to definition 3.

Definition 3 (Descent Direction) ***h*** *is a descent direction for F at x if $h^T F'(x) < 0$.*

3.1.1 The steepest Descent method

The Steepest Descent method or gradient method arises from 3.14 and satisfies:

$$\lim_{\alpha \to 0} \frac{F(x) - F(x + \alpha h)}{\alpha \parallel h \parallel} = -\frac{1}{\parallel h \parallel} h^T F'(x) = - \parallel F'(x) \parallel \cos\theta, \tag{3.15}$$

where θ is the angle between h and F'. It has been proven that the highest gain rate is succeeded when $\theta = \pi$. Thus, the steepest descent direction is given by:

$$h_{sd} = -F'(x). \tag{3.16}$$

Supposing that we have a descent direction h, we want to move to a local minimum with step αh. But we need to know how far we should go. Thus, two methods of finding the appropriate α are presented in next sections (Line search and Trust Region methods).

3.1.2 Line Search

Assume that we are in a point x and we want to move from x to x' in direction h. To find how far we should go, we study the variation of the objective function as a function of (α).

$$\phi(\alpha) = F(x + \alpha h) \tag{3.17}$$

Knowing that h is a descent direction we have, $\phi'(0) = h^T F'(x) < 0$ and equivalently $\phi(\alpha) < \phi(0)$. When α is very small, $F(x) - F(x + \alpha h)$ is small too. Thus, α must be increased. If α is too large then $\phi(\alpha) \geq \phi(0)$. Hence we have to decrease α for the descent condition to be satisfied. Finally, if α is close to the minimizer we accept the current α.

The stopping criterion for the line search at k_{th} iterate is: $|\phi'(\alpha_k)| \leq \tau |\phi'(0)|$, where τ is a small positive number. As an iterative process, line search, can be very computationally expensive. So, a stricter version of the descending condition is used. For example:

$$\phi(\alpha_k) \leq \phi(0) + \gamma_1 \phi'(0)\alpha, \qquad 0 < \gamma_1 < 1. \tag{3.18}$$

This, prevents α from being too large. Moreover, another control for avoiding α becoming too small is:

$$\phi'(\alpha_k) \geq \gamma_2 \phi'(0), \qquad 0 < \gamma_1 < \gamma_2 < 1. \tag{3.19}$$

If α satisfies both these criteria then we accept it. Otherwise, other techniques are used [19].

3.1.3 Trust Region Methods

Suppose that we have the truncated Taylor expansion $l_1(h)$ of r, in a neighbourhood of x,

$$r(x+h) \simeq l_1(h) = r(x) + J(x)h. \tag{3.20}$$

Then, by (3.1) the objective function becomes,

$$F(x+h) \simeq L_2(h) = \frac{1}{2} l_1(h)^T l_1(h) \tag{3.21}$$

$$= \frac{1}{2} r^T r + h^T J^T r + \frac{1}{2} h^T J^T J h \tag{3.22}$$

$$= F(x) + h^T J^T r + \frac{1}{2} h^T J^T J h \tag{3.23}$$

where: $r = r(x)$ and $J = J(x)$. The gradient and the Hessian of L are:

$$L'(h) = J^T r + J^T J h, \qquad L''(h) = J^T J. \tag{3.24}$$

In a trust region method, we suppose that $L(h)$ is sufficiently accurate inside a ball with radius Δ. Thus, our problem takes the form:

$$h = h_{tr} = min_{\|h\| \leq \Delta} \{L(h)\} \tag{3.25}$$

If the descending condition $(F(x+h) < F(x))$ is satisfied , then, the current approximation of the minimizer becomes $x = x + h$. Otherwise, due to the fact that h was too large and should be reduced we update Δ analogously. To handle Δ, the gain ratio the following selection for ρ is proposed:

$$\rho = \frac{F(x) - F(x+h)}{L(0) - L(h)}. \tag{3.26}$$

It is the ratio between the actual and the predicted decrease in function value. A good strategy to update Δ is that if $\rho < \frac{1}{4}$ we decide to use smaller steps, while $\rho > \frac{3}{4}$ we may need larger steps [19].

In the sequel, three other efficient methods for problem (3.1) are presented (Gauss-Newton, Levenberg-Marquardt and Powell's Dog leg methods.)

The Gauss-Newton Method

The Gauss-Newton method is the basis of the methods that are going to be further analysed in next sections. All these methods are based on a linear approximation of

the objective function. So, for small $\| h \|$ let $(l_1(h))$ be the truncated Taylor expansion of r around x (3.20). By the definition of F in(3.1),

$$F(x+h) \simeq L_2(h) = F(x) + h^T J^T r + \frac{1}{2} h^T J^T J h. \tag{3.27}$$

The problem is to find the minimum of $L_2(h)$ which is much easier than (3.1). By setting (3.24) equal to zero, the Gauss-Newton step $(\mathrm{h_{gn}})$ can be found by solving the system:

$$(J^T J) h_{gn} = -J^T r. \tag{3.28}$$

Moreover, $(\mathrm{h_{gn}})$ is a descent direction for F. From definition 3 we have:

$$h_{gn}^T F'(x) = h_{gn}^T (J^T r) = -h_{gn} (J^T J) h_{gn} < 0. \tag{3.29}$$

Hence, since $(\mathrm{h_{gn}})$ is a direction, the next step is computed by:

$$\mathbf{x} := \mathbf{x} + \alpha h_{gn} \tag{3.30}$$

and α is found by line search. Usually, in the classical Gauss-Newton method $\alpha = 1$ is used in all steps. The method with the line search can be shown to have guaranteed convergence [19] and [20].

The Levenberg-Marquardt Method

Levenberg [21] and Marquardt [22] proposed the use of a damped version of the Gauss-Newton method. According to (3.28) the new direction $(\mathrm{h_{lm}})$ is computed by the adjustment on the system:

$$(J^T J + \mu I) h_{lm} = -J^T r, \qquad \mu \geq 0. \tag{3.31}$$

The damping parameter μ has been proposed because,

- For $\mu > 0$, the matrix $(J^T J + \mu I)$ is positive definite, ensuring that h_{lm} is a descent direction.
- When μ is large then from 3.16:

$$h_{lm} \simeq -\frac{1}{\mu} F'(x), \tag{3.32}$$

  meaning a short step in the descent direction and it is good when the current iterate is far from the solution.

- If μ is small then $h_{lm} \simeq h_{gn}$ which is a good step in the final stages of the iteration (x is close to x^*).

Hence, μ is responsible for the direction and the size of the step. The first choice of μ is related to the size of the diagonal elements of the Hessian (3.9). For example, $\mu_0 = max_i\{H_{GN_{ii}}\}$. To update μ we look at the gain ratio, ρ (3.26) . If ρ is large then, $L(h_{lm})$ behaves like $F(x + h_{lm})$ and we can decrease μ so that the next step is closer to the Gauss-Newton step. Otherwise, if ρ is small then μ should be increased and hopefully get closer to the steepest descent direction. More details for the parameter μ can be found in [20].

Powell's Dog Leg Method

This trust region method, exploits the directions of the Gauss-Newton and the steepest descend methods. Powell proposed a way of finding an approximation to the trust region direction h_{tr}. As shown previously, the Gauss-Newton step (h_{gn}) is obtained by solving the system in 3.28 and the steepest descent direction, (h_{sd}) is calculated from (3.16). Due to the fact that h_{sd} is a direction and not a step, to see how far we should move, we look at the linear model:

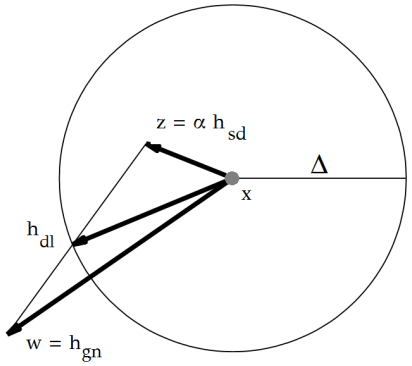


Figure 3.2: Trust region and Dog Leg step

$$r(x + \alpha h) \simeq r(x) + \alpha J(x)h \tag{3.33}$$

$$\Downarrow \tag{3.34}$$

$$F(x + \alpha h) \simeq \frac{1}{2} \parallel r(x) + \alpha J(x)h \parallel^2 \tag{3.35}$$

$$= F(x) + \alpha h^T J^T(x) + \frac{1}{2}\alpha^2 \parallel J(x)h \parallel^2 \tag{3.36}$$

It is clearly that $F(x + \alpha h)$ has minimum for

$$\alpha = -\frac{h^T J(x)^T r(x)}{\parallel J(x)h \parallel^2} = \frac{\parallel g \parallel^2}{\parallel J(x)g \parallel^2}, \quad g = J^T(x)r \tag{3.37}$$

By calculating h_{gn} and h_{sd} and supposing that we are in the current point x, we define the candidates **z** = αh_{sd} and **w** = h_{gn} for the next iteration. In this trust region method of radius Δ, Powell suggested the strategy in (Algorithm 1) and the last case is shown in (figure 3.2).

if $\| h_{gn} \| \leq \Delta$ **then**
  $h_{dl} = h_{gn}$
else if $\| \alpha h_{sd} \| \geq \Delta$ **then**
  $h_{dl} = (\Delta / \| h_{sd} \|) h_{sd}$
else
  $h_{dl} = \alpha h_{sd} + \beta(h_{gn} - \alpha h_{sd})$
  with β chosen so that $\| h_{dl} \| = \Delta$
end

Algorithm 1: Powell's Dog Leg strategy

In order to find β so that $\| h_{dl} \| = \Delta$ and with $y = z^T(w - z)$, we define the second degree polynomial $p(\beta)$:

$$p(\beta) = \| z + \beta(w - z) \|^2 - \Delta^2 = \| w - z \|^2 \beta^2 + 2y\beta + \| z \|^2 - \Delta^2 \tag{3.38}$$

In the interval $[-\infty, 0]$ exists at least one root because when $\beta \to -\infty$ then $p(\beta) \to \infty$ and $p(0) = \| z \|^2 - \Delta^2 < 0$. Furthermore in $[0, 1]$ exists another root because $p(1) = \| h_{gn} \|^2 - \Delta^2 > 0$. Due to the fact that the degree of the polynomial is two, we conclude that it has exactly two roots and we need the positive one, which can be calculated with basic algebra.

$k \leftarrow 0$; $x \leftarrow x_0$; $\Delta \leftarrow \Delta_0$; $g \leftarrow J(x)r(x)$;
$stop \leftarrow \| r(x) \|_\infty \leq \epsilon_2$ **or** $\| g(x) \|_\infty \leq \epsilon_1$
while ***not*** $stop$ ***and*** $k < k_{max}$ **do**
  $k \leftarrow k + 1$; Obtain α from (3.37)
  $h_{sd} \leftarrow -\alpha g$; Solve $J(x)h_{gn} = -r(x)$
  Obtain h_{dl} from (algorithm 1)
  if $\| h_{dl} \| \leq \epsilon_2(\| x \| + \epsilon_2)$ **then**
    $stop \leftarrow true$
  else
    $x_{new} \leftarrow x + h_{dl}$
    $\rho = \frac{F(x) - F(x_{new})}{L(0) - L(h_{dl})}$
    if $\rho > 0$ **then**
      $x \leftarrow x_{new}$; $g \leftarrow J(x)r(x)$
      $stop \leftarrow \| r(x) \|_\infty \leq \epsilon_2$ **or** $\| g(x) \|_\infty \leq \epsilon_1$
    if $\rho > 0.75$ **then**
      $\Delta \leftarrow max\{\Delta, 3 * \| h_{dl} \|\}$
    else if $\rho < 0.25$ **then**
      $\Delta \leftarrow \Delta/2$; $stop \leftarrow \Delta \leq \epsilon_1(\| x \| + \epsilon_1)$
  end
end

Algorithm 2: Powell's Dog Leg Method

Again, Powell's method as a trust region method is controlled by the gain ratio ρ at (3.26). In the previous method, (L-M) we used the gain ratio to control the damping parameter μ. In Powell's method we use ρ to control the radius Δ of the trust region. Generally, if ρ is large then the linear approximation of the objective function is good.

In this case, we can increase Δ and make larger steps that behave like Gauss-Newton direction. In contrast, when ρ is small, we reduce the radius Δ and making smaller steps that behave like the steepest descent method. Typical ρ threshold values can be found in ([19], [20], [23]).

With ϵ_1, ϵ_2 sufficiently small, a starting point x_0, an initial trust region radius Δ_0 and a control k_{max} for the iterations, a pseudo-algorithm of Powell's Dog leg Method is presented above in (algorithm 2).

Trust region algorithms have as a stopping criterion, $F'(x^*) = g(x^*) = 0$. Equivalently, for $\epsilon > 0$ and small, the stopping criterion is $\parallel F'(x) \parallel_\infty \leq \epsilon$.

3.1.4 Constrained Optimization

In the previous section, methods for the non-linear least squares problem were analysed in order to optimize the given parameters. However, in practice many parameters have natural bounds that limit their values. In the current section, the adjustment of a trust region method for use on bound-constrained problems will be discussed [24]. The dominant trust region method that will be adjusted, with the Active Set strategy in [25], is Powell's Dog Leg method. For that purpose, some preliminary tools are needed.

- **Preliminaries**

Given the lower (**l**) and upper (**u**) bounds, with $l, u \in \mathbb{R}^m$, the projection of $\mathbf{x} \in \mathbb{R}^m$ onto the feasible set X is:

$$P_X(x) = \min_{v \in X} \parallel v - x \parallel = max\{min\{x, u\}, l\} \tag{3.39}$$

The constrained set X is defined by:

$$X = \{\mathbf{x} \in \mathbb{R}^m | l_i \leq x_i \leq u_i, \quad i = 1, ..., m\} \tag{3.40}$$

A set $\mathcal{A}$ for a vector x is called active and is given by:

$$\mathcal{A}(x) = \{i | x_i = l_i \quad or \quad x_i = u_i\} \tag{3.41}$$

The inactive set, $\mathcal{I}(x)$, is the complementary set of $\mathcal{A}(x)$. Thus for the index $i \in \mathcal{I}(x)$, x_i are the inactive variables. In analogy with the theory of the non-linear least squares, the gradient of a vector must be defined. Therefore, in the constrained settings, the projected gradient of a function $r(x)$ at $x \in X$, is defined by:

$$r'(x) = \begin{cases} \frac{\partial r(x)}{\partial x_i} & \text{if } i \in \mathcal{I}(x) \quad or \quad (i \in \mathcal{A}(x) \quad and \quad \frac{\partial r(x)}{\partial x_i} < 0) \\ 0 & \text{otherwise} \end{cases} \tag{3.42}$$

Additionally, the reduced Hessian $(H^R_{i,j})$ cannot be passed over.

$$H^R_{i,j} = \begin{cases} \frac{\partial^2 r(x)}{\partial x_i \partial x_j} & \text{if } i \in \mathcal{I}(x) \quad or \quad j \in \mathcal{I}(x) \\ \delta_{ij} & \text{otherwise} \end{cases} \tag{3.43}$$

3.1.5 The Modified Dog Leg Algorithm (MDL)

In order to make the Dog Leg algorithm more robust we introduce the Active Set strategy. That is, the linear model $L_1(h)$ of the objective function, the term $J^T J$ must be replaced with $H^R_{i,j} = D_{\mathcal{I}}(x_k) J^T J D_{\mathcal{I}}(x_k) + D_{\mathcal{A}}(x_k)$.

Where, $D_{\mathcal{I}}$ is the diagonal matrix with,

$$[D_{\mathcal{I}}(x)]_{jj} \begin{cases} 1, j \in \mathcal{I}(x) \\ 0, j \in \mathcal{A}(x) \end{cases} \qquad \text{and} \quad D_{\mathcal{A}}(x) = I - D_{\mathcal{I}}(x) \tag{3.44}$$

As a further improvement, we replace the gradient of r in (3.37) with the constrained version (3.42). Finally, at k_{th} iteration, if x_k is outside of the constraints after the $k-th$ Dog-Leg solution, we project the new step x_k onto the constrained set X as:

$$x_k = \mathcal{P}_{\mathcal{X}}(x_k + h) - h \tag{3.45}$$

In the sequel, examples for the applications of the (MDL) algorithm for the estimation of the IVIM parameters are shown.

3.2 Results - Limitations

3.2.1 Data Description

To begin with, for the needs of our study, two middle-aged patients (**P1** and **P2**) with no organ relevant pathology were selected. For each patient, an expert radiologist selected regions of interest (**ROIS**) on well perfused pancreatic and liver tissues, avoiding areas of severe contamination of motion artifacts. Diffusion weighted images are obtained from a 1.5 Tesla Siemens MR scanner with b-values of [0 50 100 150 200 500 1000 1500 2000 3000]. The exponentially decaying signal intensities of the diffusion weighted images as a function of the b-values for each patient and for each anatomical area, are shown in the following figures. ROIS are marked with red color.

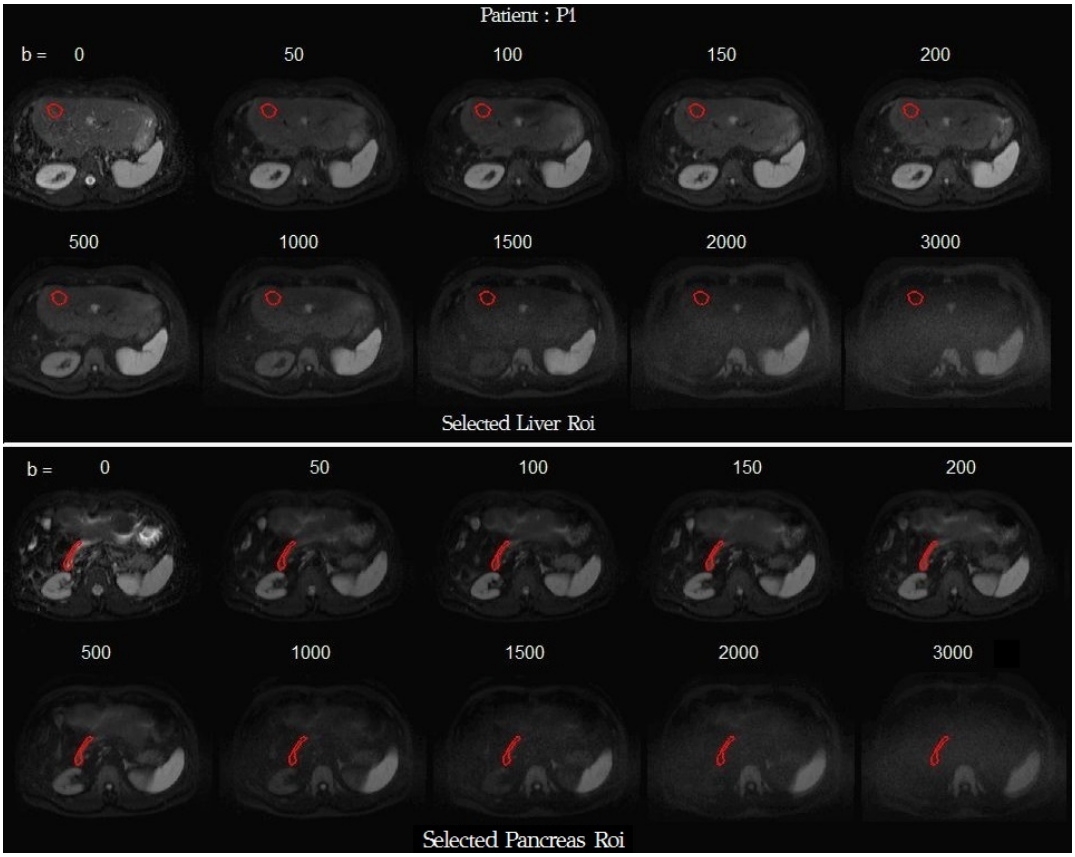


Figure 3.3: Selected ROIS for patient P1

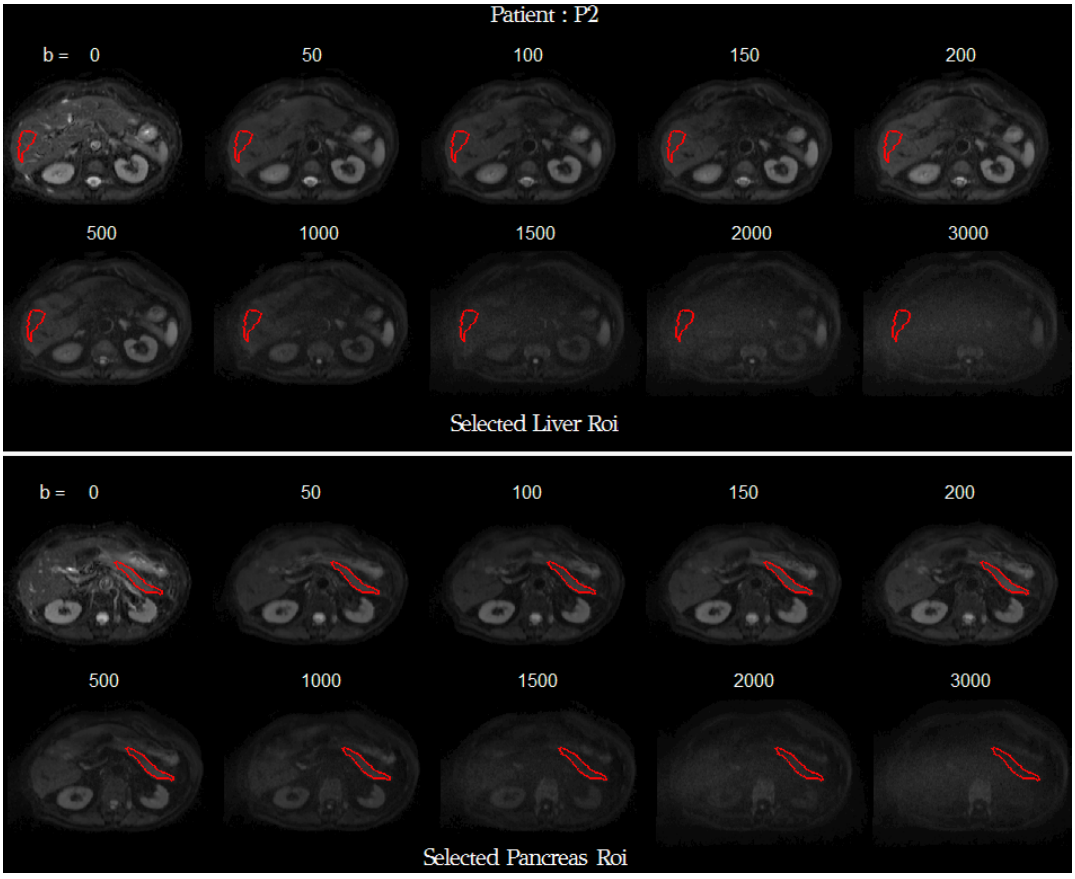


Figure 3.4: Selected ROIS for patient P2

3.2.2 Stating the optimization problem

The main goal of the DWI analysis and specifically of the studied IVIM model (section 2.5.6), is to produce parametric maps (bio-markers) of the examined area of an organ that highlight water diffusivity and micro-perfusion activity. These maps are mainly calculated from the non-linear least squares theory as it is described in section 3.1.

Technically speaking, for each pixel in a ROI of a specific anatomical area the main goal is to determine the values (D, D^*, f) from the IVIM model as given by (2.34). Non-linear least squares are applied to the exponentially decaying signal intensity curve of each pixel. As defined by the IVIM model, the decaying curve is fitted to the DWI data and the IVIM parameters are calculated.

As can be seen in figure 3.5, for each DW-image in respect to its b-value, a point inside a ROI is selected and an indicative exponentially decaying curve of a pixel in the ROI as a function of the b-value is depicted in figure 3.6. (Signal intensities are presented with $(*)$)

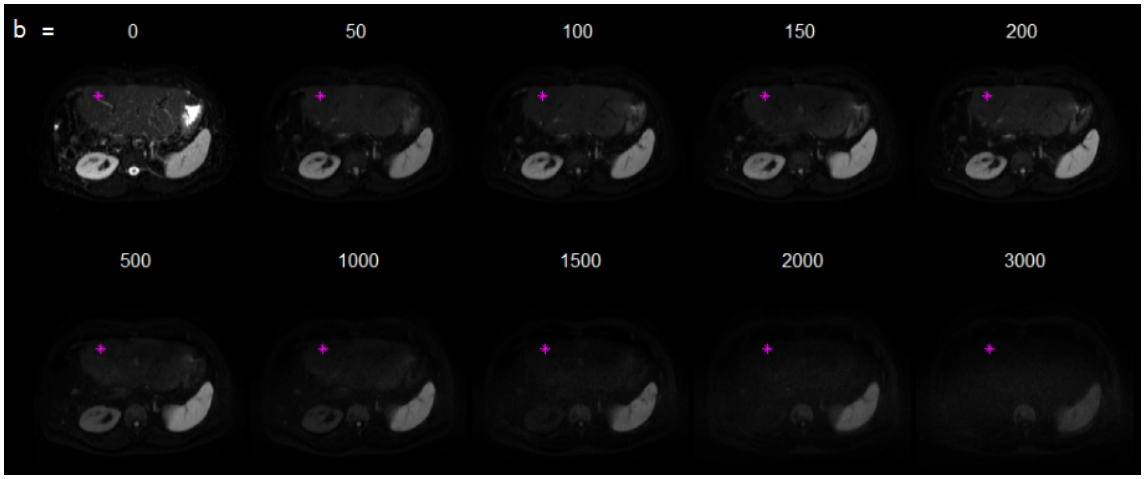


Figure 3.5: Signal intensities for a pixel from multiple b-values inside a ROI

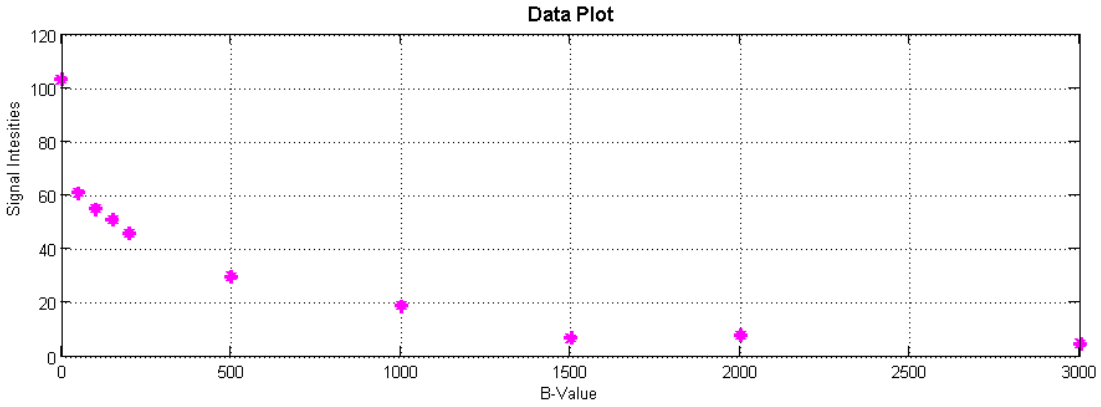


Figure 3.6: Signal intensities for a pixel in b-value space

The main aim of this process is to fit the DWI data $(b_i, y_i) \quad i = 1, ..., 10$ to the model function G. To be more comprehensive, the model function according to the IVIM

theory is given by:

$$G(\mathbf{x}, b) = S_0((1 - x_1)e^{-bx_2} + x_1 e^{-bx_3}) - S_b \tag{3.46}$$

For each pixel in the ROI, S_0 stands for the the signal intensity with no diffusion gradient ($b - value = 0$). The vector $\mathbf{x} = [x_1, x_2, x_3] = [f, D, D^*]$ of (2.34) and is calculated via the optimization procedure of the non-linear least squares. Thus, the objective function F to minimize is:

$$F(x) = \frac{1}{2}\sum_{i=1}^{n} r_i^2(x) \quad \text{with} \quad r_i\,(x) = y_i - G(\mathbf{x}, b), \quad i = 1, ..., 10 \tag{3.47}$$

In bibliography, the estimated IVIM parameters lie on the following ranges: the perfusion fraction $f \in [0, 1]$, the diffusion coefficient $D \in [0.2, 3] \times 10^{-3}$ [9] and the pseudo-diffusion coefficient $D^* \simeq 10 \times D$ [18]. This fact led us to use the modified version of Dog-Leg algorithm presented in section 3.1.5.

By using the constrained version of the Dog-Leg algorithm, for a random pixel inside the ROI, with initial guess $x_0 = [0.1, 10^{-3}, 30 \times 10^{-3}]$, tolerance $\epsilon = 10^{-9}$, $k_{max} = 100$ iterations, the lower $\mathbf{l} = [0.01, 0.2 \times 10^{-3}, 10 \times 10^{-3}]$ and the upper $\mathbf{u} = [0.9, 4 \times 10^{-3}, 300 \times 10^{-3}]$ bounds, algorithm converged to: $\mathbf{x} = [0.475, 1.18 \times 10^{-3}, 2.96 \times 10^{-2}]$ a clinically relevant result. The fitted model function can be seen in (figure 3.7).

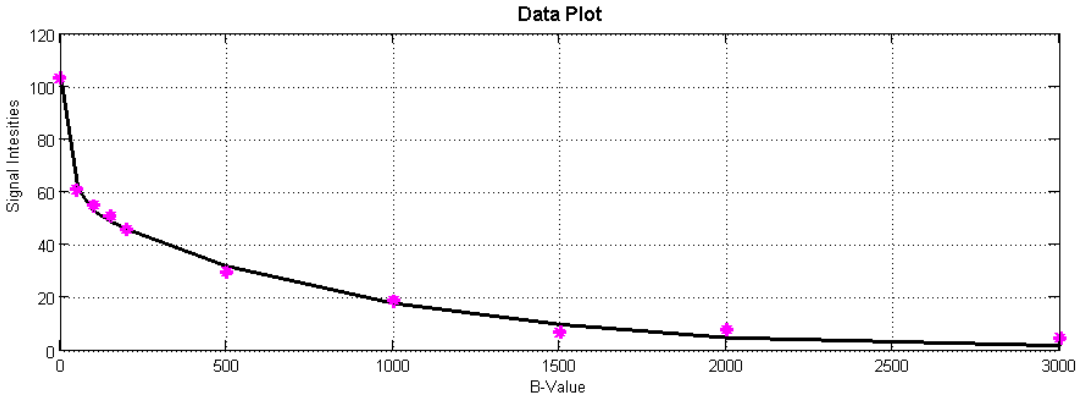


Figure 3.7: Fitted data on the objective function

The above fitting procedure is repeated for every pixel inside the selected ROI. Before presenting the results for each anatomical area and for each patient, a metric for the validity of our results is needed.

In statistics, the parameter R-squared (R^2) is commonly used as a measure of similarity between the data and the model function G. $R^2 \equiv 1 - \dfrac{SS_{res}}{SS_{tot}}$ where: SS_{res} is the sum of squares of residuals, also called the residual sum of squares $SS_{res} = \sum_i (y_i - G_i)^2$ and SS_{tot} the total sum of squares (proportional to the variance of the data), $SS_{tot} = \sum_i (y_i - \bar{y})^2$ where $\bar{y}$ is the mean of the observed data: $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. $(R^2) \in [0, 1]$

and 1 stands for perfect similarity, while 0 for dissimilarity. This measure is suitable for models that have only one variable that needs optimization. In this work, three variables needed to be optimized and a proposed measure for similarity is the adjusted R-squared ($\overline{R}^2$) [26]. In analogy with R-squared, the adjusted R-squared takes values from -1 to 1 where 1 is the indicator of perfect similarity. $\overline{R}^2 = 1-(1-R^2)\frac{n-1}{n-p-1}$ where p is the total number of the optimized variables in the model function, and n is the sample size.

3.2.3 Limitations

This sections's purpose is to present the results (when the IVIM model is applied to the liver and the pancreas) of using the non-linear least squares.

To begin with, the modified Dog Leg algorithm with initial guess $x_0 = [0.1, 10^{-3}, 30 \times 10^{-3}]$, lower $\mathbf{l} = [0.01, 0.2 \times 10^{-3}, 10 \times 10^{-3}]$ and upper $\mathbf{u} = [0.9, 4 \times 10^{-3}, 300 \times 10^{-3}]$ bounds, is used to depict the IVIM parameters. First, f, D, D^* and $\overline{R}^2$ maps are presented for patient's P1 liver (figure 3.8 and pancreas 3.9). After that, the same maps are presented for patient's P2 liver and pancreas in figures 3.10 and 3.11 respectively.

Also, for every pixel in the selected ROI after each minimization, the adjusted-R-square is calculated. Hence, an ($\overline{R}^2$) map is produced to describe how accurately the model function is fitted to the data. Furthermore for every map, the histogram of each parameter is calculated and shown in the following figures.

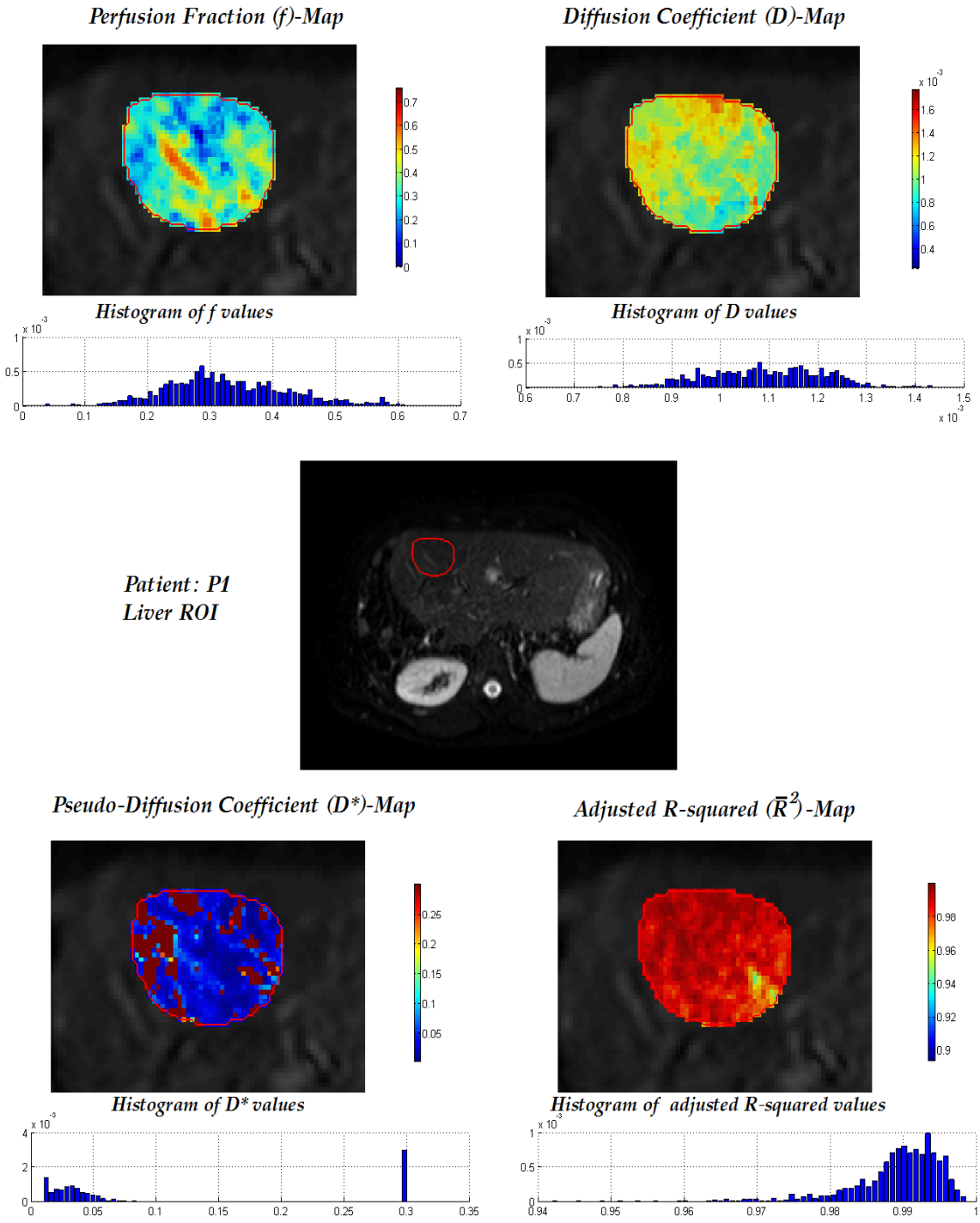


Figure 3.8: Patient: P1, liver Maps and histogram of values for each map

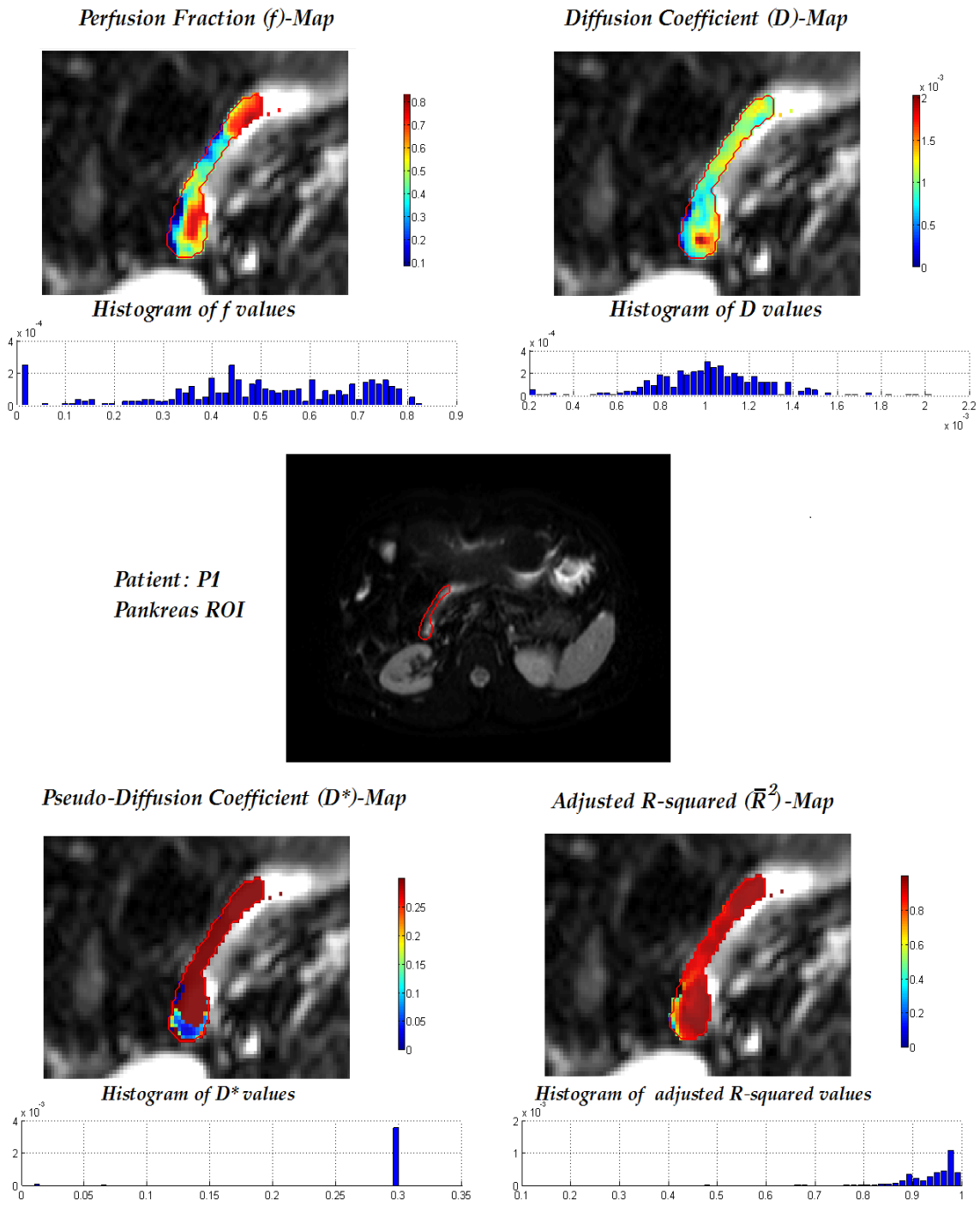


Figure 3.9: Patient: P1, Pancreas Maps and histogram of values for each map

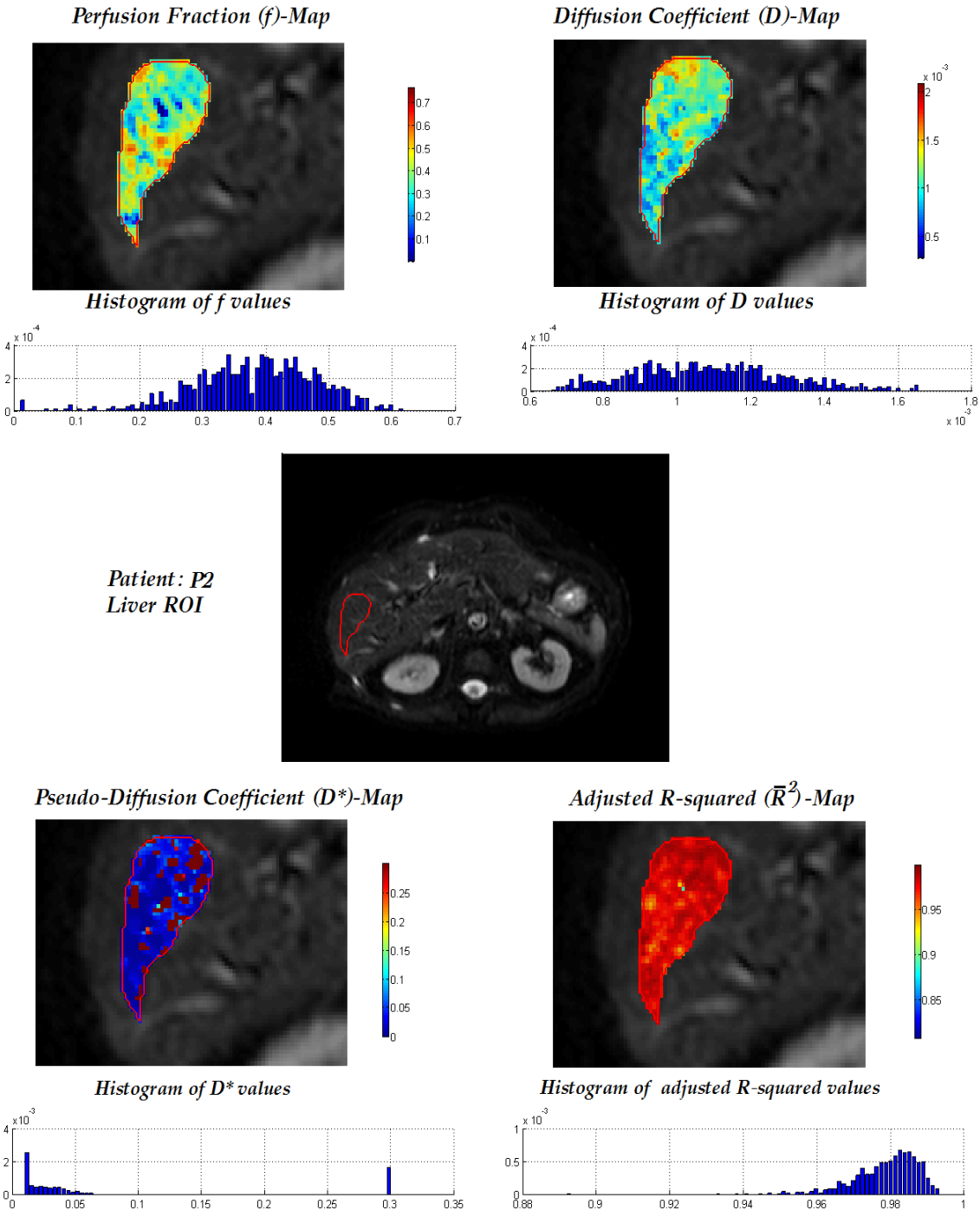


Figure 3.10: Patient: P2, liver Maps and histogram of values for each map

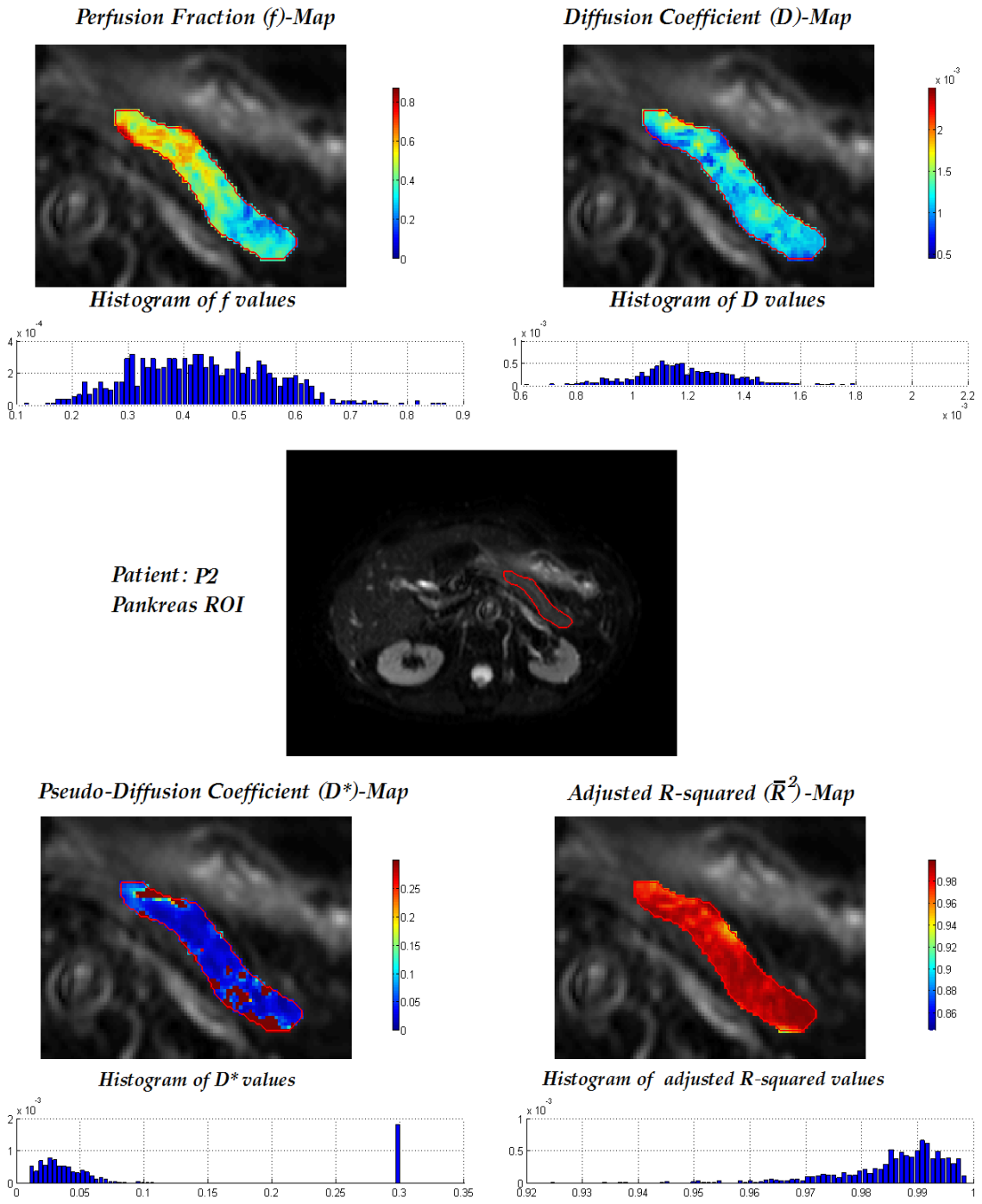


Figure 3.11: Patient: P2, Pancreas Maps and histogram of values for each map

Previously, the derived from the IVIM model parameters, f, D, D^* and $(\overline{R}^2)$ maps have been presented. The $(\overline{R}^2)$ was big enough in every case, ensuring that the fitting of the DWI data was accurate. From the statistical point of view the fitting process seemed to have worked properly and accurately. In practice, a duo tone in all D^* maps has been observed. In every D^* map, a great percentage of the parameter was on the initial upper bound (0.3) (figure 3.12). The reason why that happens is, because being

in exponential form and being one order of magnitude bigger than D, e^{D^*} is affected significantly. Thus, changes in D^* take us out of bounds and from the active sets strategy (3.45) the projection of the last step solution, onto the constrained set (3.40) take us to the nearest bound.

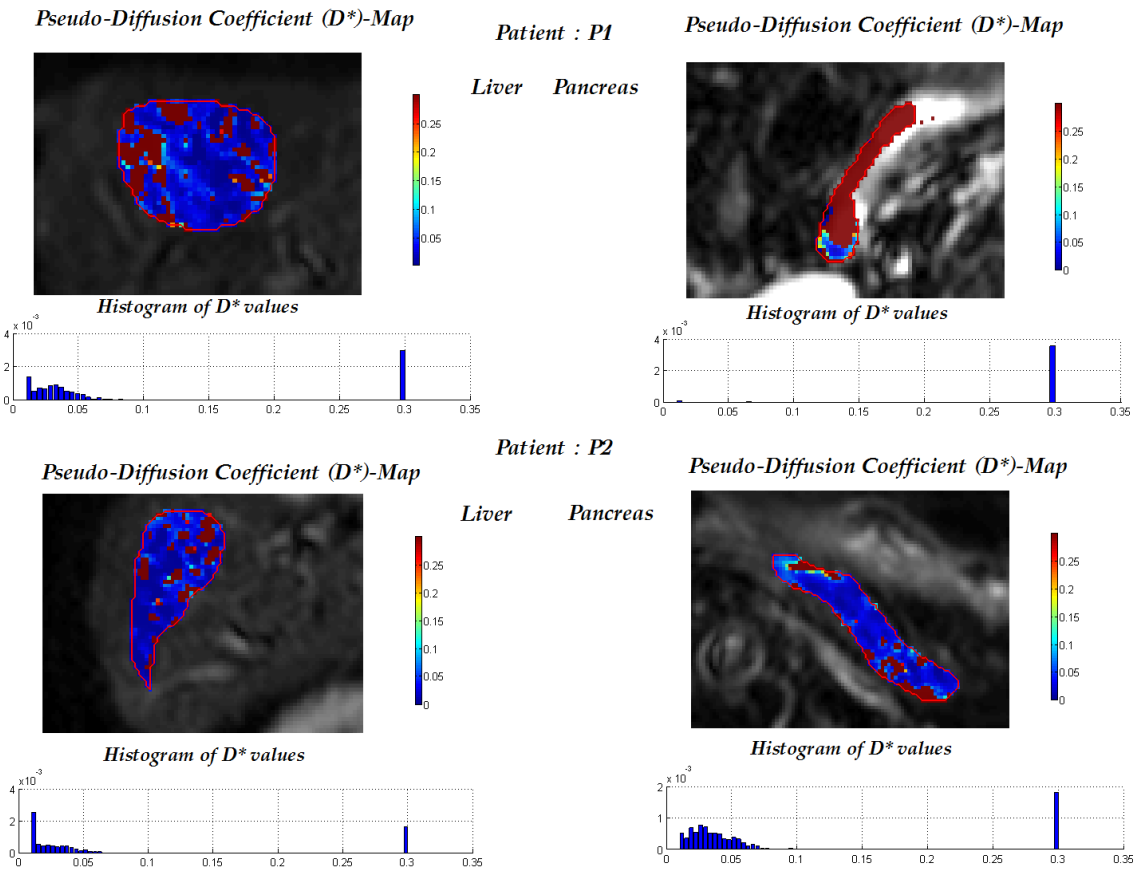


Figure 3.12: The problem with D* parameter

4

Proposed Method

Non-linear least squares, as an iterative process, are highly dependant on the initial starting point. By altering the initial point, the algorithm may take us to different result. Thus, reliability is not succeeded. Data fitting is a complex procedure, especially when more than one variable need to be optimized. The more the optimization parameters, the most local minima and maxima, thus the decision of which solution is accepted, becomes more difficult. In order to control the result, constrained optimization techniques are used with strict constraints. If the result is out of the preferred bounds then the final result is projected to the nearest bound.

In order to simplify the objective function, many published research works, propose partial fitting [27], [28]: the diffusion coefficient (D) is obtained after regression to the simple ADC model 2.19. After that, with D known, the IVIM model is fitted to data with unknown parameters only D^* and f.

To resolve this problem, an automated way of finding suitable initial guesses for our fitting algorithm is proposed. Hence, for every pixel in the selected ROI, the optimization is performed by using **three** variables f, D and D^*. In the following paragraphs, the way of finding these initial guesses will be analysed.

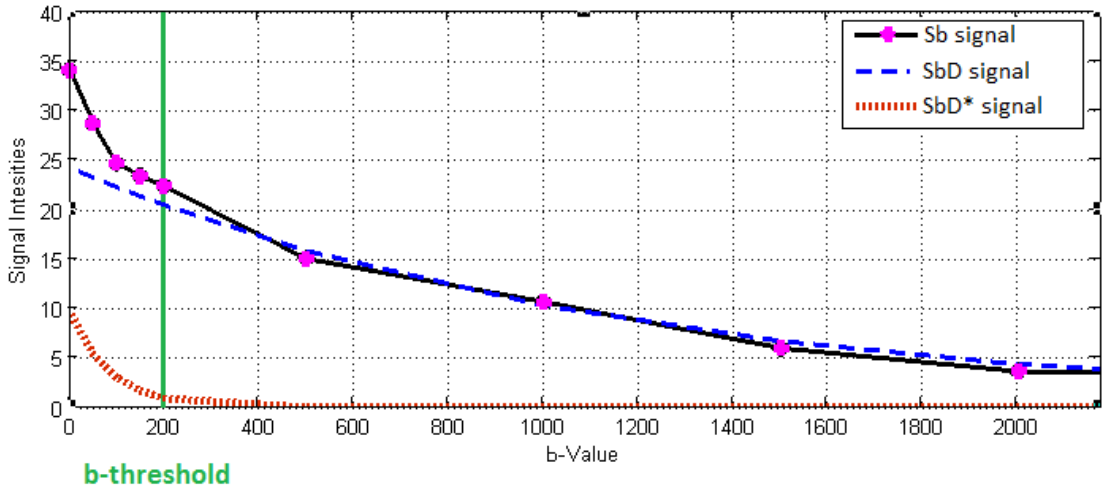


Figure 4.1: Purple dots represent the MR signal. Blue line stands for the mono-exponential diffusion signal and red line for the pseudo-diffusion signal. Green line is the critical b-value

First of all, it is known that for b-values greater than 200 sec/mm^2, the micro-perfusion effect is eliminated [28]. Therefore, this critical value will be referred as **b-threshold**.

Thus, the contributed to the IVIM micro-perfusion effect signal $\frac{S_{b_{D^*}}}{Sb_0} = fe^{-bD^*}$ fades out quickly before the **b-threshold** value. Furthermore, the primal signal S_b becomes equal to the mono- exponential signal of diffusion $\frac{S_{bD}}{Sb_0} = (1-f)e^{-bD}$ after the **b-threshold** value. The signals above are presented in (figure 4.1).

Having only the signal from the MR scanner S_b, the procedure of calculating the initial guesses before the fitting algorithm is described below.

Calculation of initial D (D_0)

For b-values $b_i >$ **b-threshold**, the mono-exponential diffusion model is used to determine D, taking into account all possible combinations $\{i, j\}$ of b-values concluding to the median of D.

$$D = \frac{1}{(b_i - b_j)} ln\left(\frac{Sb_j}{Sb_i}\right) \tag{4.1}$$

Calculation of initial f (f_0)

For the combinations of b_i previously described, is assumed that the micro-perfusion effect is eliminated, thus $S_b = S_{bD}$. So, with known D, f is a result of :

$$\frac{S_b}{Sb_0} = (1-f)e^{-bD} \tag{4.2}$$

Calculation of initial D* (D_0^*)

Now we have D and f known. For b-values $b_i <$ **b-threshold** the micro-perfusion effect is present. In order to calculate D^* we simulate $S_{b_{D^*}}$ as: $S_{b_{iD^*}} = S_{b_i} - S_{b_iD}$. Finally D^* is calculated from:

$$\frac{S_{ib_{D^*}}}{Sb_0} = fe^{-b_iD_i^*} \quad \text{for all i} \tag{4.3}$$

and by taking the median of D^* we conclude to the final initial value from the IVIM model.

Finally, by this initialization we have, [initial_f, initial_D, initial_D*]. In order to fit the IVIM function to the data with the constrained version of Dog-Leg algorithm, the lower and upper bounds are needed. To define the bounds, we reasonably set the lower bound **l** to **z** percent lower than the initial _values (**l** = initial_values - **z** % initial_values) and the upper bound **u** to be **z** percent higher than the initial_values (**u** = initial_values + **z** % initial_values) .

Thus, a **self adjusted initialization** as a pre-phase of the fitting algorithm have been proposed. In the next chapter, the comparisons between results obtained by the proposed method with the self adjusted initialization and those of the standard method, for the estimation of the IVIM parameters is presented.

4.1 Results

With our **initial_values**, **b-threshold** = 200 and **z** = 30, the derived from the IVIM model maps for patient's P1 liver and pancreas are shown in (figures 4.2 and 4.3) respectively. Moreover the same maps are shown for patient's P2 liver (figure 4.4) and pancreas (figure 4.5).

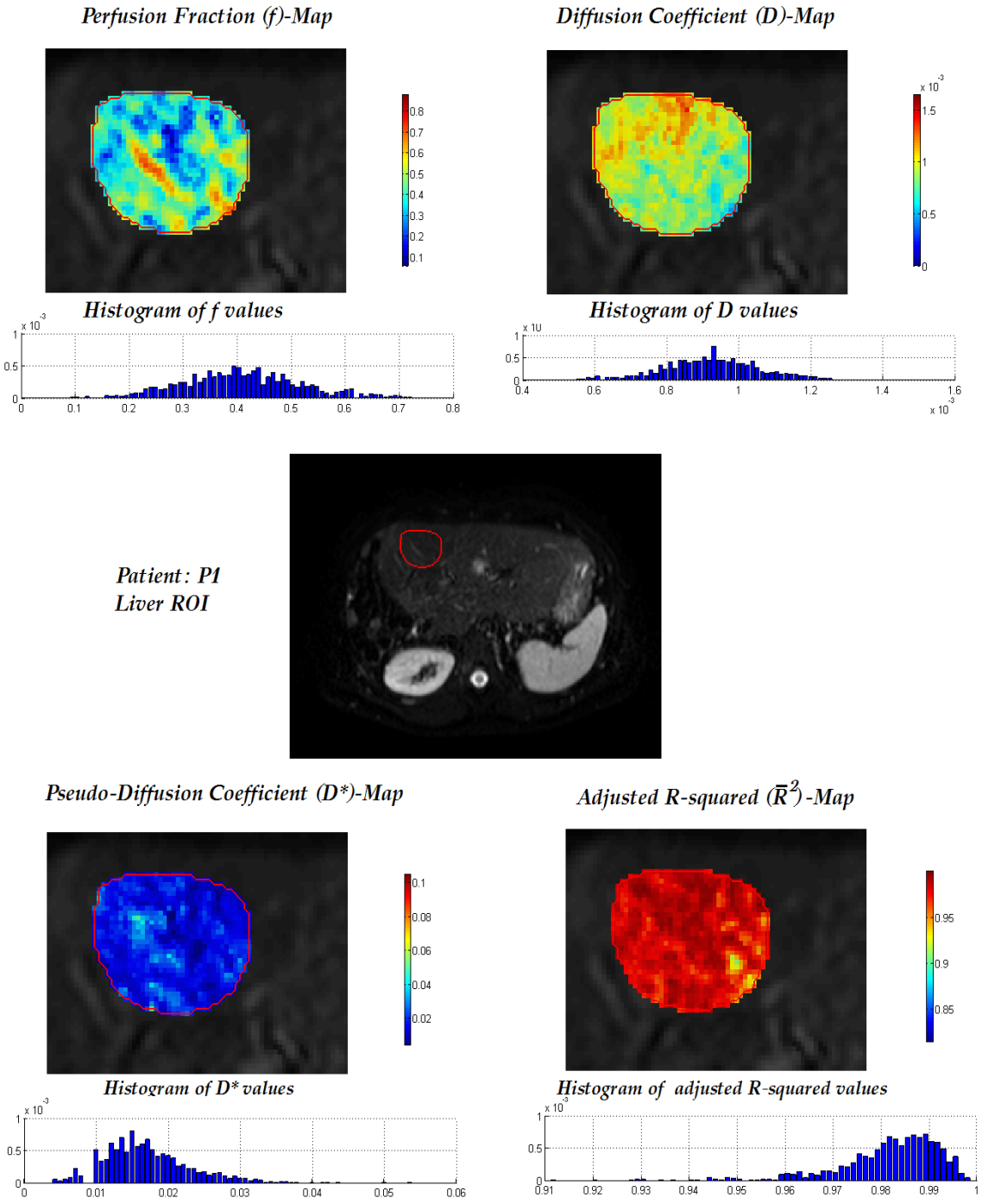


Figure 4.2: Patient: P1, liver Maps and histogram of values for each map

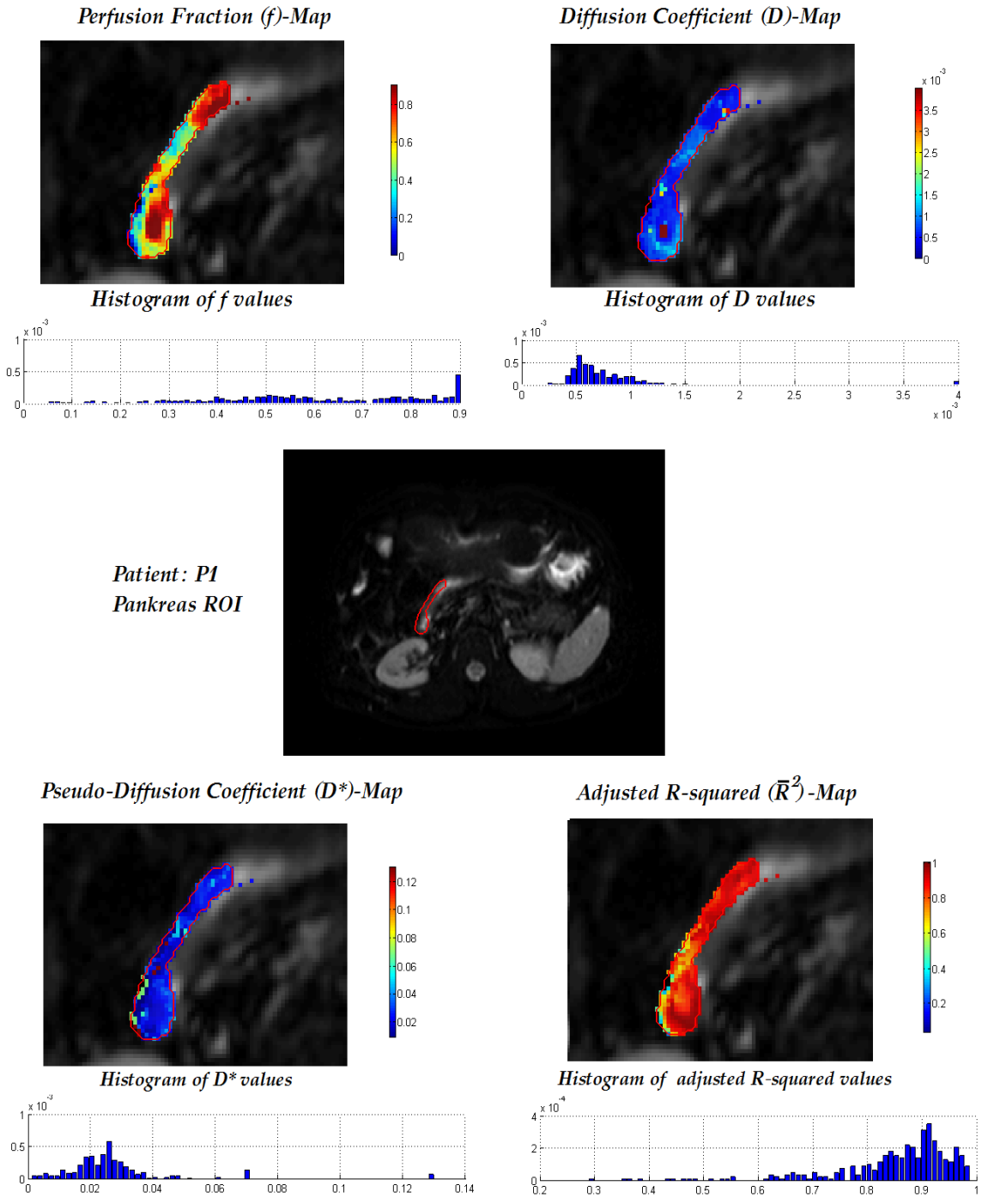


Figure 4.3: Patient: P1, Pancreas Maps and histogram of values for each map

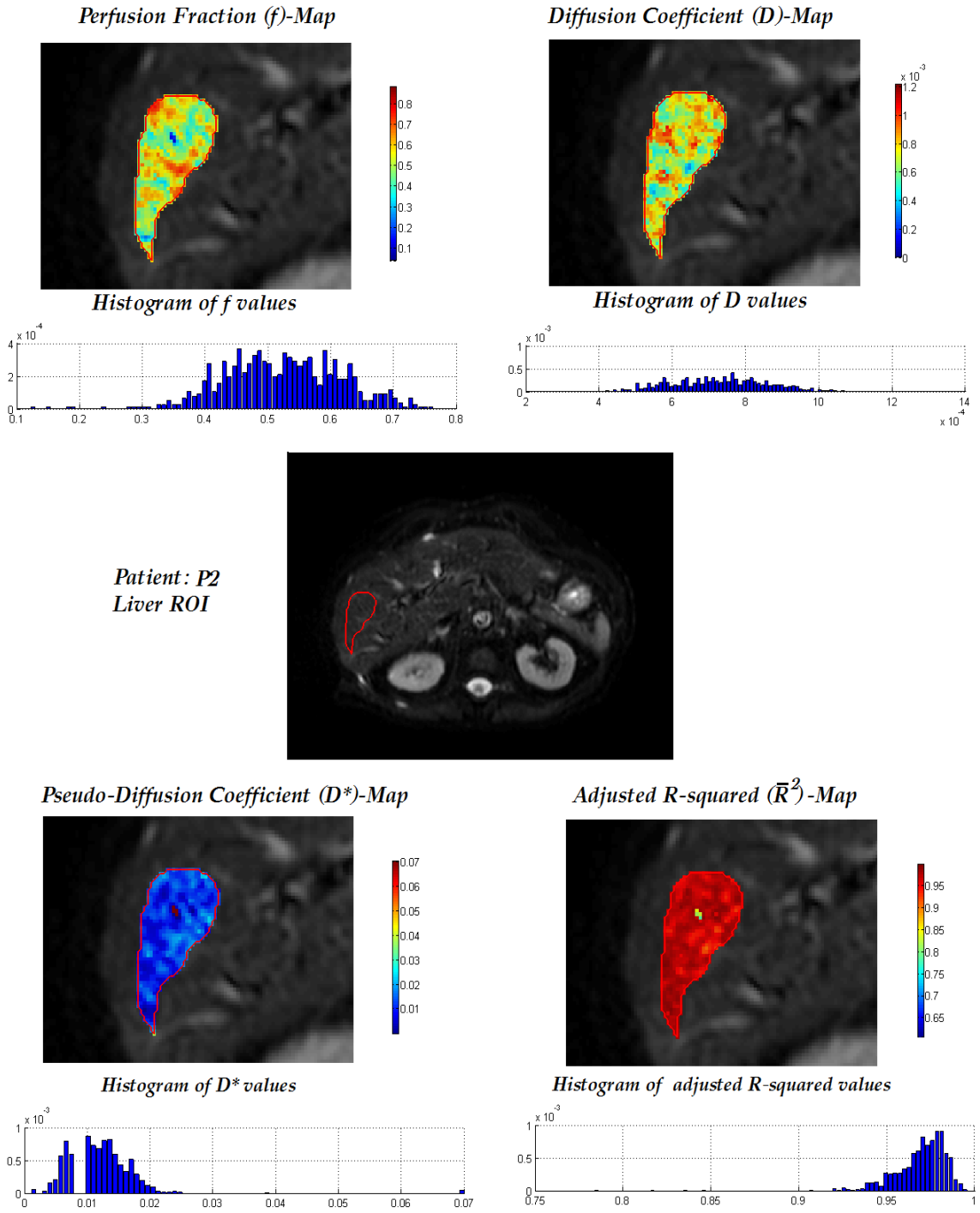


Figure 4.4: Patient: P2, liver Maps and histogram of values for each map

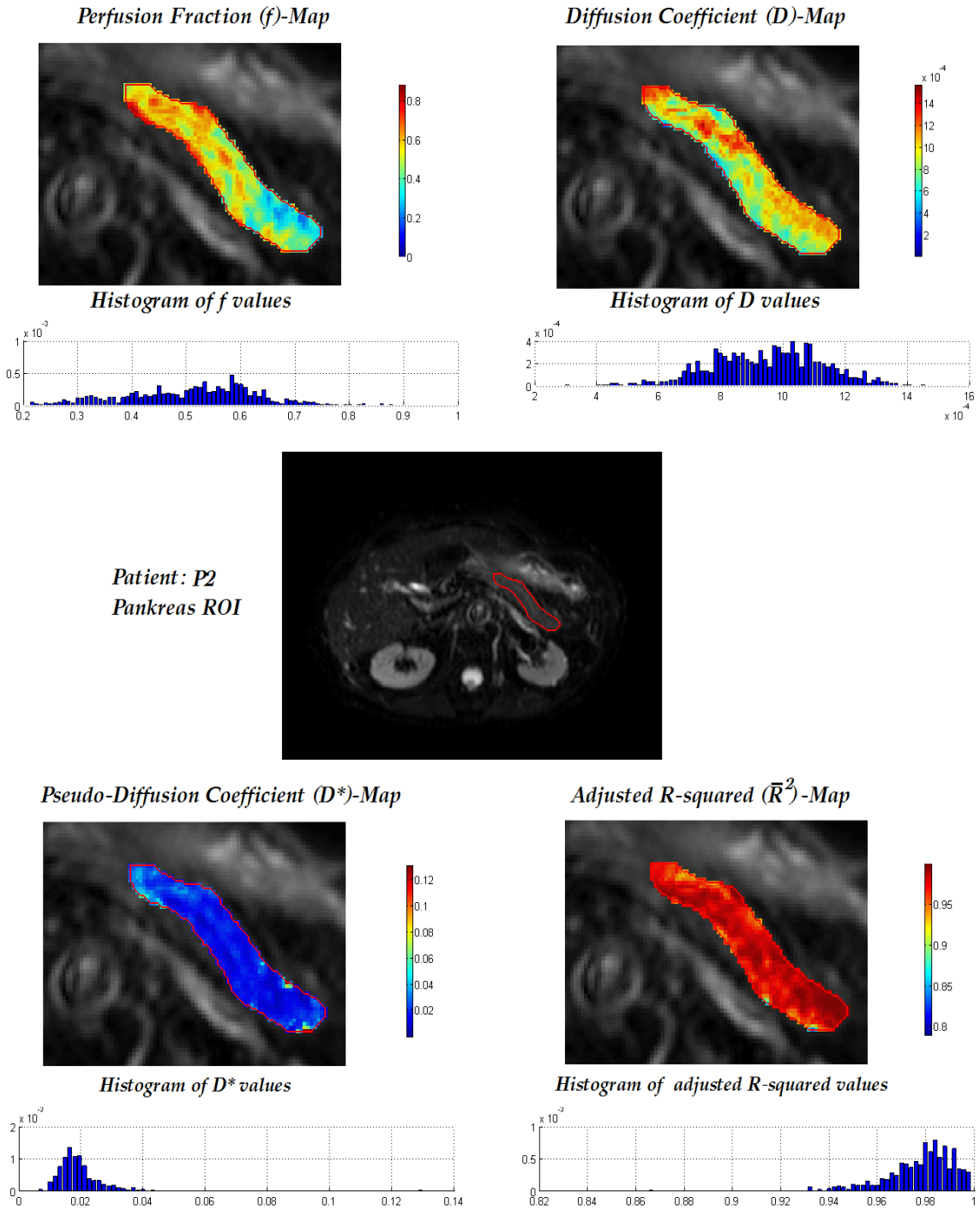


Figure 4.5: Patient: P2, Pancreas Maps and histogram of values for each map

4.2 Comparison-Discussion

To sum up, data fitting is a complex procedure, especially when more than one variables need to be optimized. The more the optimization parameters, the most local minima and maxima, thus the decision of which solution is accepted, becomes more difficult. In this work, the results after fitting were clinically irrelevant, since the D^* parameter was in the upper bound in a great percentage.

To resolve this problem, suitable initial guesses and personalized bounds, for our fitting algorithm, were built according to the IVIM theory. Thus, the optimization is succeeded by the use of three variables (f, D and D^*). The IVIM parametric maps were presented as well as , the histogram of the values of each parameter. For each parameter, the histogram tends to take the shape of the Gaussian distribution. Our similarity measure adjusted R-squared ($\overline{R}^2$) has indicated that the fitting was accurate with values greater than 0.91 approximately. Hence, the problem with the upper bound is overtaken by our initialization.

A comparison of the D^* maps before and after the personalized bounds is shown in (figure 4.6) for patient P1 and in (figure 4.7) for patient P2.

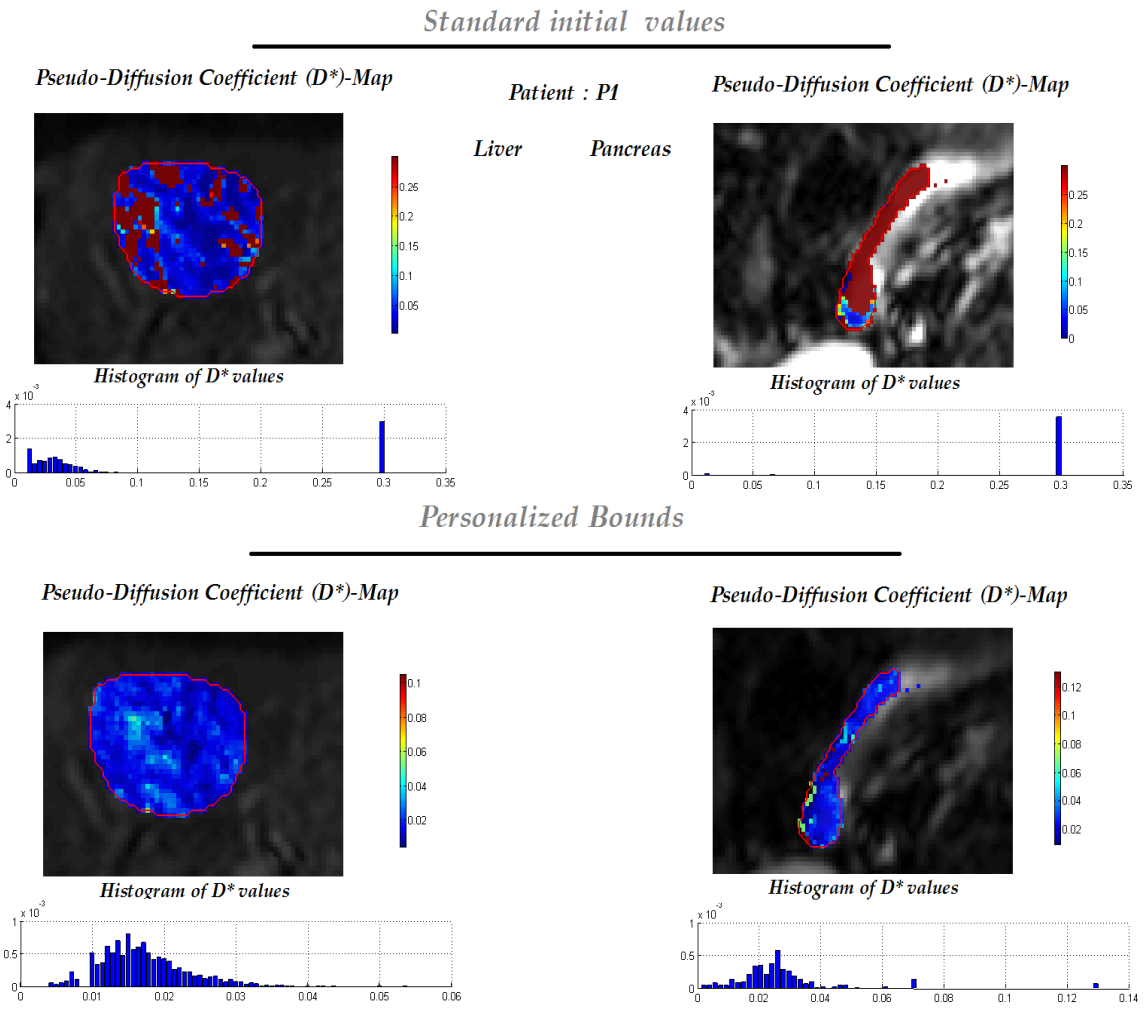


Figure 4.6: Comparison between D* values from two different methods of patient: P1.

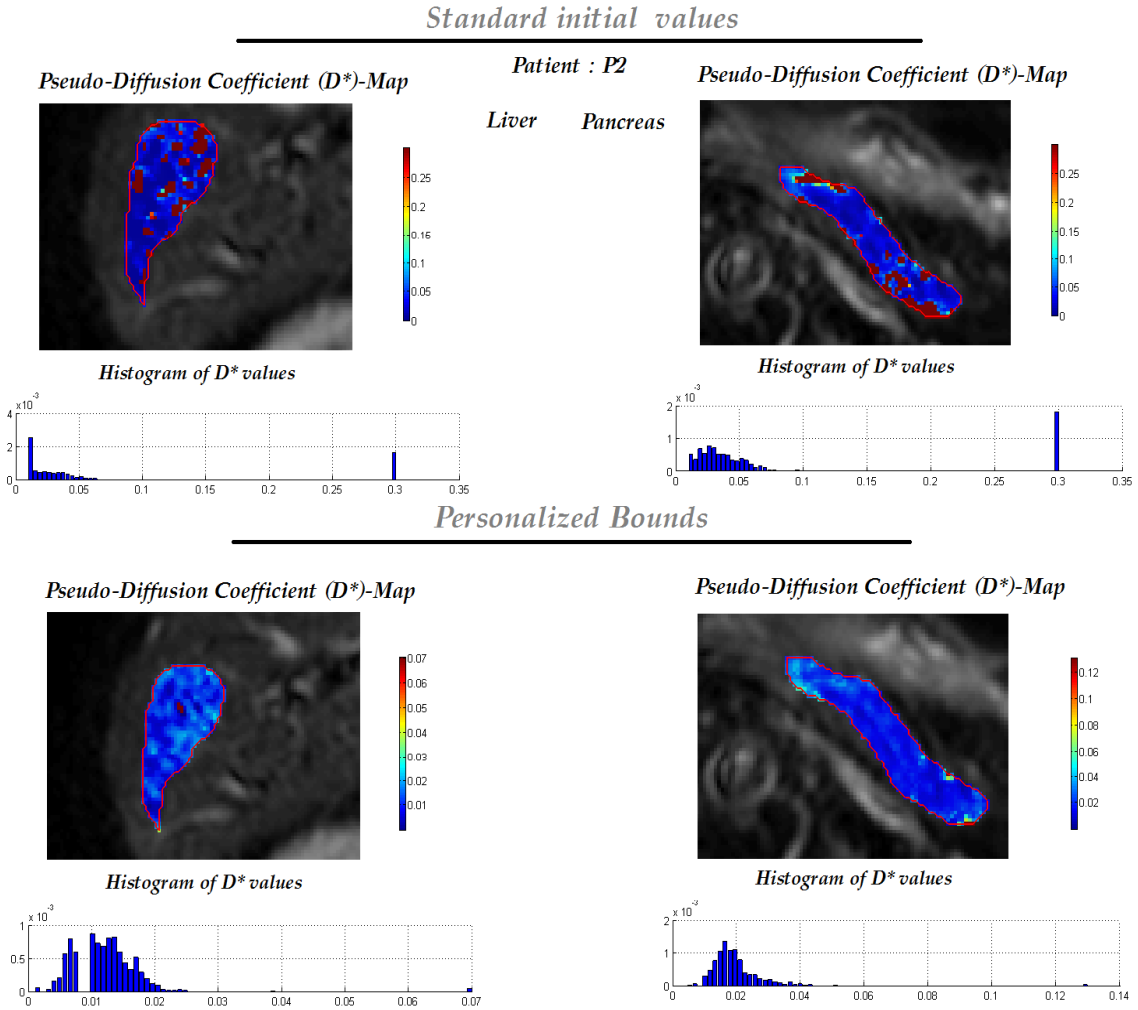


Figure 4.7: Comparison between D* values from two different methods of patient: P2.

5

Conclusion

In this master thesis, the basic physics behind MRI is introduced and a thorough study of the diffusion analysis is given in order to depict parametric maps through the the DWI. These maps, give quantitative information about the tissue water diffusivity, assisting doctors and radiologists in giving clinical answers about several pathologies to anatomical areas of the human body.

Intravoxel Incoherent Motion (IVIM) model and its provided parametric maps, produce information about water diffusivity and tissue vascularity in organs such as the liver and the pancreas. The dominant tool for the construction of the maps, is non-linear least squares. The result of the non-linear least squares, as an iterative process, is unreliable because the method is highly dependant on the initial starting point.

However, a mathematical framework as a pre-processing phase of the fitting process has been developed. After comparison between the standard fitting procedure and the self adjusted initialization method, more accurate and clinically accepted quantitative parametric maps for (f, D and D^*) have been obtained with the proposed method introduced in this thesis.

5.1 Future Work

To examine how reliable the proposed method is, further analysis on altering the **z** percentage of initial values that affect the bounds as proposed in chapter 4 must be done. Furthermore, the optimal selection of the critical **b-threshold** value must be further studied. Also, a statistical analysis of the proposed method to a great number of data, could ensure the reliability of the method.

From another perspective, with the advances in the MR hardware technology, it is possible to acquire higher b-values (greater than 1000 s/mm^2) and thus it is possible to observe deviation from the Gaussian approximation of the diffusion process (kurtosis phenomenon). More complex models are necessary in order to optimize fitting when MR acquisition includes higher b-values. In that case, equations will be adjusted to account for both micro-circulation of blood in the capillary network and kurtosis. The proposed model for IVIM-kurtosis would demand accurate fitting of four parameters rather than three in the present study. Thus, a similar approach as in the IVIM model will be considered to the IVIM-kurtosis model.

Bibliography

[1] Catherine Westbrook. *MRI at a Glance, 2nd ed.* Blackwell Science Ltd, 2002.

[2] D.M. Koh and Harriet C Thoeny. *Diffusion weighted MR Imaging, Applications in the Body*. Springer Berlin Heidelberg, 2010.

[3] Ashkan A Malayeri, Riham H El Khouli, Atif Zaheer, Michael A Jacobs, Celia P Corona-Villalobos, Ihab R Kamel, and Katarzyna J Macura. Principles and applications of diffusion-weighted imaging in cancer detection, staging, and treatment follow-up. *Radiographics : a review publication of the Radiological Society of North America, Inc*, 31(6):1773--91, October 2011.

[4] David G. Brock Elliott L. Mancall. *GRAY'S Clinical Neuroanatomy The Anatomical Basis of Clinical Practice*. Elsevier, 2011.

[5] Marinus T. Vlaardingerbroek. *Magnetic Resonance Imaging Theory and Practice*. Springer-Verlag Berlin Heidelberg, 2003.

[6] Dominik Weishaupt, Victor D. Köchli, and Borut Marincek. *How does MRI work?* Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[7] R. Nick Bryan. *Introduction to the Science of Medical Imaging*. 2010.

[8] Paul Tofts. *Quantitative MRI of the Brain*. John Wiley & Sons Ltd, 2003.

[9] Patric Hagmann, Lisa Jonasson, Philippe Maeder, Jean-philippe Thiran, Van J Wedeen, and Reto Meuli. CENTRAL NERVOUS SYSTEM : STATE OF THE ART Understanding Diffusion MR Imaging Techniques : From Scalar Imaging to Diffusion. *RadioGraphics*, pages 205--224, 2006.

[10] H. C. Torrey. Bloch Equations with Diffusion Terms. *Physical Review*, 104(3):563--565, 1956.

[11] Breton Eric Denis Le Bihan. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161:401--407, 1986.

[12] D Le Bihan, J F Mangin, C Poupon, C a Clark, S Pappata, N Molko, and H Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of magnetic resonance imaging : JMRI*, 13(4):534--46, April 2001.

[13] C-F Westin and S E Maier. A Dual Tensor Basis Solution to the Stejskal-Tanner Equations for DT-MRI. *Proc. Intl. Soc. Mag. Reson. Med*, 10:39335, 2002.

[14] Andrew L Alexander, Jee Eun Lee, Mariana Lazar, and Aaron S Field. Diffusion Tensor Imaging of the Brain. *NIH Public Access*, 4(3):316--329, 2008.

[15] P Mukherjee, J I Berman, S W Chung, C P Hess, and R G Henry. Diffusion tensor MR imaging and fiber tractography: theoretic underpinnings. *AJNR. American journal of neuroradiology*, 29(4):632--41, April 2008.

[16] Alexandre Vignaud, D Ph, Alain Luciani, and Alain Rahmouni. Intravoxel Incoherent Motion (IVIM)*: A Potential Application in Cirrhosis Assessment. *Magnetom Flash*, (Ivim):36--39, 2009.

[17] Denis Lallemand Le Bihan, Denis, Eric Breton. Seperation of Diffusion and Perfusion in Intravoxel Incoherent Motion MR Imaging. *Radiology*, 168:497--505, 1988.

[18] R. L. Cooper, D. B. Chang, a. C. Young, C. J. Martin, and D. Ancker-Johnson. Restricted diffusion in biophysical systems: experiment. *Biophysical Journal*, 14(3):161--177, 1974.

[19] Poul Erik Frandsen, Kristian Jonasson, Hans Bruun Nielsen, and Ole Tingleff. *Unconstrained optimization*. 1999.

[20] K Madsen, H B Nielsen, and O Tingleff. *Methods for non-linear least squares problems*. 2004.

[21] K Levenberg. A method for the solution of certain non-linear problems in least squares. *The Quarterly of Applied Mathematics*, II(2):164--168, 1944.

[22] Donald W. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431--441, 1963.

[23] Nobuko Sagara and Masao Fukushima. A hybrid method for the nonlinear least squares problem with simple bounds. *Journal of Computational and Applied Mathematics*, 36(2):149--157, 1991.

[24] Dimitri P. Bertsekas. Projected Newton Methods for Optimization Problems with Simple Constraints. *SIAM Journal on Control and Optimization*, 20(2):221--246, 1982.

[25] J N Holt and R Fletcher. An Algorithm for Constrained Non-linear Least-squares. *Journal of the Institute of Mathematics and its Applications*, 23:449--463, 1979.

[26] Bg Tabachnick and Ls Fidell. *Using multivariate statistics*. Pearson Education, 6th edition, 2007.

[27] Christian Federau, Kieran O'Brien, Reto Meuli, Patric Hagmann, and Philippe Maeder. Measuring brain perfusion with intravoxel incoherent motion (IVIM): Initial clinical experience. *Journal of Magnetic Resonance Imaging*, 39(3):624--632, 2014.

[28] Yu-Chuan Hu, Lin-Feng Yan, Lang Wu, Pang Du, Bao-Ying Chen, Liang Wang, Shu-Mei Wang, Yu Han, Qiang Tian, Ying Yu, Tian-Yong Xu, Wen Wang, and Guang-Bin Cui. Intravoxel incoherent motion diffusion-weighted MR imaging of gliomas: efficacy in preoperative grading. *Scientific Reports*, 4(16):7208, 2014.