HAL
open science

Signal and Noise in Diffusion Magnetic Resonance Images

Emmanuel Caruyer

To cite this version:

Emmanuel Caruyer. Signal and Noise in Diffusion Magnetic Resonance Images. Medical Imaging. 2007. dumas-00880381

HAL Id: dumas-00880381

https://dumas.ccsd.cnrs.fr/dumas-00880381v1

Submitted on 5 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Signal and Noise in Diffusion Magnetic Resonance Images

Emmanuel Caruyer, MSc in Computer Science

June 11, 2007

Supervisor: Rachid Deriche

Odyssée Project-Team, INRIA Sophia-Antipolis.

Abstract

This report summarizes the work I did within the Odyssée research group, in INRIA Sophia-Antipolis, under the supervision of Rachid Deriche. We have been working on problems related to noise in medical images, more specifically in diffusion weighted MRI, originating from physical process we are able to model. In the light of these models, the purpose of our work was to evaluate existing reconstruction methods, and to propose some refinement.

I want to greatly thank Rachid firstly for the topic he proposed, which I realized is widely of interest, for the discussions we had during my internship and finally for his contribution to making my everyday life at INRIA pleasant.

Contents

1 Physics of diffusion MRI **5**
1.1 Nuclear magnetic resonance . . . 5
1.2 Diffusion weighted images . . . 6
1.3 Noise in magnitude MRI . . . 7
1.3.1 Probability density function of the magnitude data . . . 8
1.3.2 Rician noise and high SNR . . . 8

2 Reconstruction techniques **10**
2.1 The model of Gaussian diffusion . . . 10
2.2 Linear reconstruction techniques . . . 11
2.2.1 Linear Least squares . . . 11
2.2.2 Weighted least squares . . . 11
2.2.3 Iterated Least Squares . . . 12
2.3 Non-linear tensor estimation . . . 13
2.3.1 Non-linear least squares . . . 13
2.3.2 Maximum likelihood estimation . . . 14

3 Comparison and statistics on reconstruction techniques **18**
3.1 Methods . . . 18
3.2 LS and Weighted LS . . . 19
3.3 Maximum-likelihood estimation . . . 20

4 Conclusion and discussion **23**

Introduction

Diffusion Magnetic Resonance Imaging (MRI) has recently appeared as a new imaging technique, providing a unique tool to investigate *in vivo* the fine tissue structure, especially within white matter in either the brain or the spinal chord. Specifically, it consists in measuring the decay in MR signal produced by the diffusion of water molecules when magnetic field gradients (called diffusion gradients) are applied during the imaging sequence. Given that the movement of water molecules is constrained by the underlying white matter fibers, one expects to retrieve information about the microscopic tissue structure from diffusion measurements.

Diffusion imaging raises new issues for MR signal processing. Indeed, the diffusion weighted images are acquired in such a way that the signal to noise ratio decreases dramatically. As reviewed before by Sijbers *et al.* [1], MR images with a very poor SNR are to be processed carefully. Moreover, recent studies by Basu *et al.* [2] and Fillard *et al.* [3] show that the noise characteristics of MR signal may introduce a bias in water diffusion parameters estimation if an inadequate method is used for estimation.

The aim of this work was to perfom an in depth, quantitative study of both the SNR and the estimation method on the accuracy of the reconstruction. As an introduction to the reconstruction techniques presented in our report, we briefly present in the first chapter the core concepts of diffusion MRI, focusing on the noise properties. Then the second chapter deals with reconstruction techniques, while the last chapter shows the results of our estimation methods on synthetic dataset.

Chapter 1

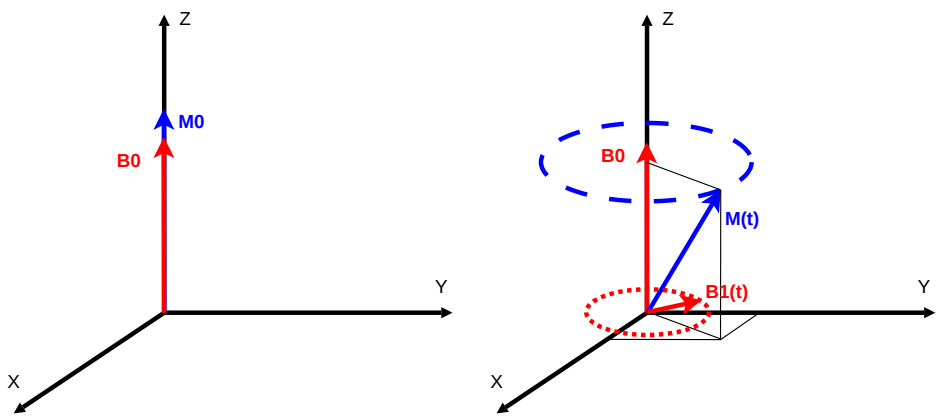
Physics of diffusion MRI

Magnetic resonance imaging allows to indirectly measure the density of protons contained in water molecules, so that various tissues contrast with each other in the images. We first recall in this chapter the basics of nuclear magnetic resonnance, and show how it can be sensitized to proton diffusion; then we discuss the particularities of this kind of imaging, and show the signal processing issues it involves.

1.1 Nuclear magnetic resonance

The core property of nuclei such as hydrogen is their intrinsic magnetic moment, so that when plunged in a strong magnetic field $\mathbf{B_0}$ they tend to align with it. Thus within a sample containing water for instance, a magnetic moment $\mathbf{M_0}$ appears at the macroscopic scale.

Moreover the microscopic moment of each proton precesses with a given frequency (the so-called Larmor frequency) around this field, and under the effect of an electromagnetic wave tuned at the Larmor frequency $\omega_0 = \gamma\|\mathbf{B_0}\|$, they precess in phase: a resonance phenomenon appears, as depicted in figure 1.1. Depending on the shape and the length of the $\mathbf{B_1}$-pulse, it is possible either to cancel (90°-pulse) or to negate (180°-pulse) the z-component of the magnetic moment. The measurement of electromagnetic field induced by relaxation after applying the rotating magnetic field $\mathbf{B_1}$ (called the Free Induction Decay) is the core idea of magnetic resonance imaging.



(a) Equilibrium

(b) Pertubation by rotating field $\mathbf{B_1}$

Figure 1.1: Macroscopic response to the perturbation by $\mathbf{B_1}$

The magnitude of the FID is directly related to the mean density of protons within the sample. Yet this principle only deals with the analysis of homogeneous samples; magnetic resonance imaging adds a means to match measured signal and spatial localisation. Indeed with an adequate acquisition sequence involving gradients of the magnetic field, it is possible to spatially modify the resonant frequency, as it is directly related to $\mathbf{B_0}$. Hence it is possible to determine the spatial contributions to the FID, via a Fourier transform. As the details of slice and line selection are beyond the scope of our work, we will not discuss them further. Instead, we propose to present the particularities of diffusion acquisition sequences.

1.2 Diffusion weighted images

As first reported in 1965 by Stejskal and Tanner [4], and experimented by Le Bihan *et al.* [5] in diffusion imaging, it is possible to apply gradients during the acquisition sequence to measure coherent motion of water molecules within a voxel. An example of acquisition sequence is shown in figure 1.2. Actually, before and after the 180°-pulse, a gradient of $\mathbf{B_0}$ is applied as a pulse of duration δ.

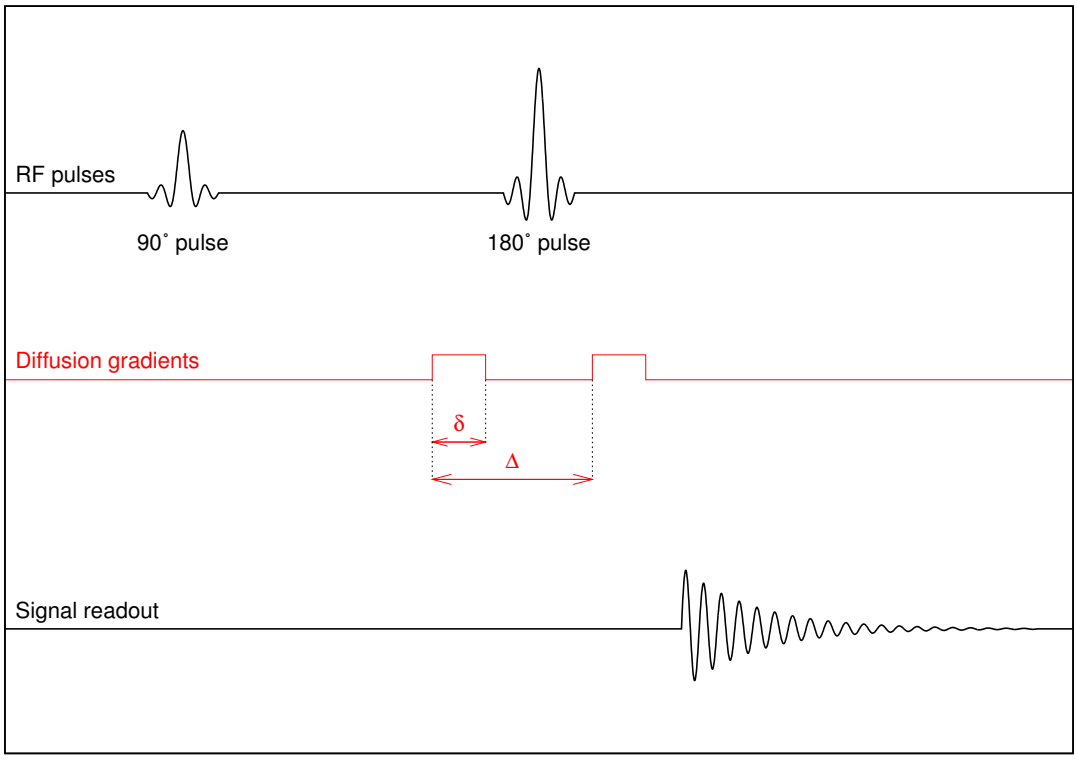


Figure 1.2: Acquisition sequence of Stejskal and Tanner

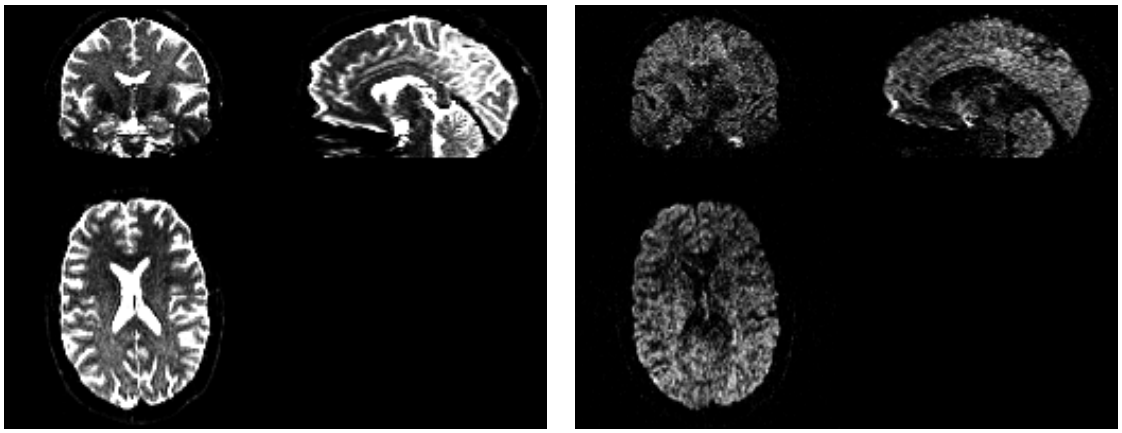
Modifying the value of the $\mathbf{B_0}$ field modifies the precessing frequency of spins, so that after the first gradient pulse is applied, the phase of spins is modified, and the phase shift depends on the position. As the 180°-pulse negates the phase, applying the same gradient pulse after it should cancel the phase shift previously introduced. However if particles have moved between the two gradient pulses, the phase shift do not compensate, and as a consequence the FID measured is attenuated in comparison to the FID without any diffusion gradient. The expected

signal S_i is given by the equation of Stejskal and Tanner:

$$S_i = S_0 e^{-bD(\mathbf{g_k})}$$

with the ponderation b-value $b = \gamma^2\delta^2(\Delta - \frac{\delta}{3})G_k^2$, where S_0 is the signal without diffusion gradient, γ is the gyromagnetic ratio, Δ and δ are represented in the acquisition sequence in figure 1.2, G_k is the strength of the magnetic field gradient and $D(\mathbf{g_k})$ characterizes the diffusion in the direction of the applied gradient. Repeating the process for various directions $\mathbf{g_k}$ offers a means to sample the properties of anisotropic diffusion.

In order to get a good angular resolution, it may be necessary to acquire a very large number of images; doing so is clinically unfeasible unless the acquisition time is reduced, which inevitably implies a reduction of the SNR. Moreover, the free induction decay measured in presence of a diffusion gradient may be strongly attenuated with respect to the signal without gradient (referred to as the $\mathbf{B_0}$ image), as witnessed by the shots of real data in figure 1.3. These two last arguments explain why diffusion weighted images (DWIs) are particularly noisy.



(a) $\mathbf{B_0}$-image (b) Diffusion weighted image

Figure 1.3: Comparison of MR image with and without diffusion gradient, b-value is 1000.

Before introducing the reconstruction models and algorithms in the next section, we now review a few properties of noise in magnitude images, which demonstrate that a particular focus is required for estimation.

1.3 Noise in magnitude MRI

As a pseudo-periodic signal, the FID can be represented by a complex number; actually it is measured through an induced electric current in two coils by the scanner. Sijbers reminds [1] that given the origins of noise, we can consider that the complex acquired data is corrupted by Rician noise. As the Fourier transform (which relates signal in the k-space to the signal in our 3D world) is a linear transform, then the signal after Fourier transform remains Gaussian distributed.

Although most scanners can measure the complex signal, it is common practice to work only with magnitude. However the operation of extracting magnitude from complex signal is non-linear. Indeed, $M = \sqrt{R^2 + I^2}$, where R and I represent the real and complex components of the complex signal, respectively.

1.3.1 Probability density function of the magnitude data

Sijbers *et al.* derive the PDF of the magnitude data, simply rewriting the complex Gaussian PDF with polar coordinates (M, ϕ), and integrating over the phase domain. The parameters of the Rician distribution are A, which is the value of interest, and σ which is the noise parameter; the PDF writes as:

$$p(M|A,\sigma) = \frac{M}{\sigma^2} \exp\left(-\frac{A^2+M^2}{2\sigma^2}\right) I_0\left(\frac{AM}{\sigma^2}\right)$$

where I_0 is the modified Bessel function of the first kind, defined as:

$$I_0(z) = \int_0^{2\pi} e^{z\cos(\phi)} d\phi$$

1.3.2 Rician noise and high SNR

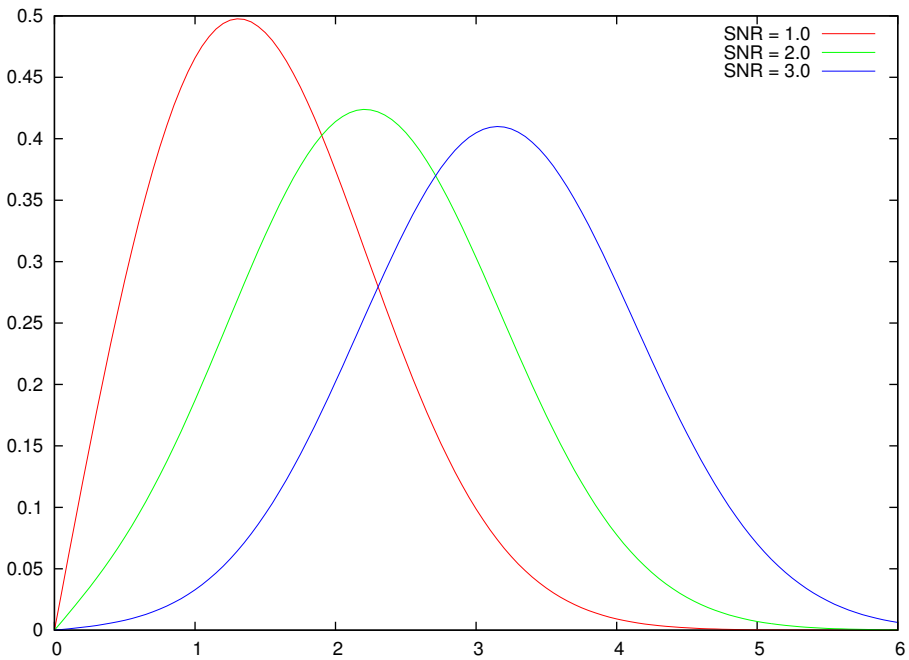


Figure 1.4: Rice PDF for various SNR values.

The figure 1.4 suggests that the Rician PDF differs from the Gaussian one for low SNR values, but seems to approximate the gaussian one when the SNR increases. Actually the first order approximation of $M = \sqrt{(A_R+\eta_R)^2 + (A_I+\eta_I)^2}$ in $(\eta_R/M, \eta_I/M)$ gives (we write $A = \sqrt{A_R^2 + A_I^2}$):

$$\begin{aligned} M &= A + \frac{A_R}{A}\eta_R + \frac{A_I}{A}\eta_I + o(\frac{\eta_R}{M}, \frac{\eta_I}{M}) \\ &= A + \eta \text{ with } \eta = \frac{A_R}{A}\eta_R + \frac{A_I}{A}\eta_I \end{aligned}$$

and as η_R and η_I are independant, and follow a normal distribution, then $\eta \sim \mathcal{N}(0, \sigma^2)$.

To have a more quantitative idea of the link between Gaussian and Rician distributions, a natural way is to use the Kullback-Leibler distance, defined as:

$$D_{KL}(p,q) = \int_{-\infty}^{+\infty} \left(p(x)\log\frac{p(x)}{q(x)} + q(x)\log\frac{q(x)}{p(x)}\right) dx$$

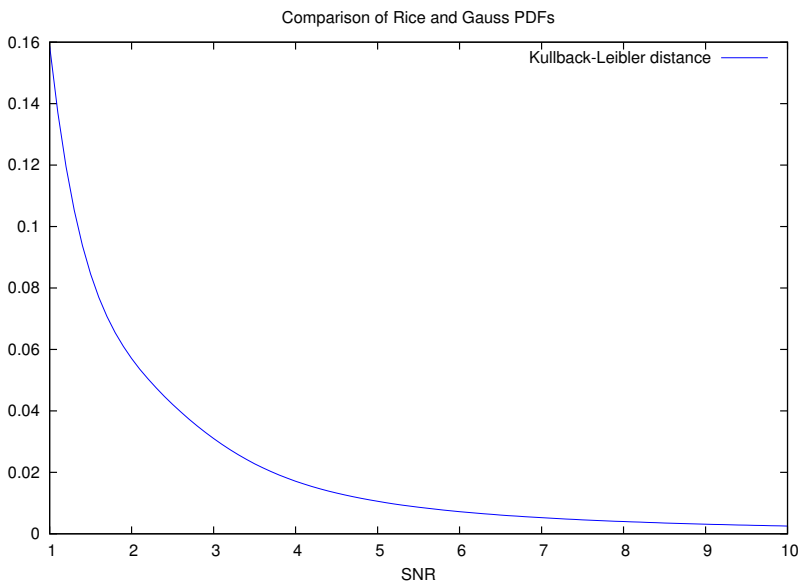


Figure 1.5: Kullbach-Leibler distance between Rician and Gaussian PDFs.

We plotted this distance against SNR in figure 1.5. The distance dramatically decreases when the SNR increases, thus for sufficiently high SNR, we can expect reconstruction techniques based on the classical assumption of Gaussian noise to give very similar results to techniques taking into account the true nature of the noise. The next two chapters present some comparative results of the estimation techniques, in order to validate this assertion, and to get an idea of what "sufficiently high" SNR means.

Chapter 2

Reconstruction techniques

In this second chapter, we first present the model of diffusion tensor, historically the first complex model introduced by Basser *et al.* [6] to describe information measured in diffusion MRI. As reviewed in the previous section, we have a good knowledge of the statistics of noise in the DWIs. Given that, we consider the maximum likelihood estimation technique, which explicitly accounts for the Rician nature of noise, and compare it to classical linear estimation techniques.

2.1 The model of Gaussian diffusion

The diffusion can be described by the probability p for a molecule to undergo a translation $\mathbf{x}$ over the time τ. It is known that for unconstrained diffusion, a Gaussian PDF is a good model to describe the diffusion. The diffusion tensor $\mathbf{D}$ is yet defined as the covariance matrix of the corresponding random variable, and assuming a Gaussian diffusion, the tensor is sufficient to describe water diffusion, as the probability density function is given by:

$$p(\mathbf{x}|\tau, \mathbf{D}) = \frac{1}{\sqrt{(4\pi\tau)^3 \det(\mathbf{D})}} \exp\left(-\frac{\mathbf{x}^T\mathbf{D}^{-1}\mathbf{x}}{4\tau}\right)$$

We remind that clinically and in research area, the main interest of diffusion MRI is to study the fine structure of the brain white matter. This part of the brain is a collection of fiber bundles, and one expects to recover the orientation of underlying fibers from diffusion measurements. In particular, in case of very anisotropic diffusion, it is very likely that the principal eigenvector of the diffusion tensor gives the direction of an underlying fiber tract.

Also, some scalars calculated from the tensor carry information about the diffusion (and so far remain one of the only application clinically in use). These are principally the fractionnal anisotropy (FA) and the trace. The FA dicriminates between cigar-shaped (FA$\simeq$1) and sphere-shaped tensors (FA$\ll$1), and thus is high in very structured regions. The trace is an index of the mean diffusivity, and is high in water-filled regions (such as ventricules in the brain). Even these scalar indices produce images where contrast non visible in structural MRI appears.

Although it is known that the tensor model is unable to describe complex tissue structure, such as fibre crossing, overlapping or kissing, we kept working with this model. Indeed we present estimation techniques with a special focus on noise properties, and wanted to keep a model simple enough.

2.2 Linear reconstruction techniques

Assuming the model of Gaussian diffusion, the equation of Stejskal and Tanner can be reformulated as:

$$S_k = S_0 e^{-b\mathbf{g_k}^T \mathbf{D} \mathbf{g_k}} + \eta_k$$

where η_k represents the uncertainty on the raw measurements.

2.2.1 Linear Least squares

The equation above can be linarized, we have then:

$$-\frac{1}{b} \ln\left(\frac{S_i}{S_0}\right) = g_k^T \mathbf{D} g_k + \epsilon_k$$

where ϵ represents the error on the transformed signal. For convenience, we use the notations introduced by Hasan *et al.* [7]: the 6 independant elements of the tensor $\mathbf{D}$ (it is a covariance matrix, so it is symmetric) are placed into a vector $\mathbf{d} = [D_{xx}\ D_{yy}\ D_{zz}\ D_{xy}\ D_{xz}\ D_{yz}]^T$, and the directions $\mathbf{g_k}$ are reformulated as vectors of 6 components: $\mathbf{h_k} = [g_x^2\ g_y^2\ g_z^2\ 2g_xg_y\ 2g_xg_z\ 2g_yg_z]^T$, in turn placed into a matrix $\mathbf{H} = [\mathbf{h_k}]_k$. Finally the vector $\mathbf{y}$ of log-measurements is introduced: $\mathbf{y} = [-1/b\ \ln(S_k/S_0)]_k$. The reformulation is simply:

$$\mathbf{y} = \mathbf{H}\mathbf{d} + \epsilon$$

Hence given the data records for the N directions $\mathbf{g_k},\ k = 1 \ldots N$, provided that we have more equations than unknowns: $N \geq 6$, we have an overconstrained linear system, which can be solved using classical Least of Squares (LS) criterion minimization. Note that for the particular case $N = 6$, Rathore [8] pointed out that there exists some singular gradient schemes, so that they are non-collinear, but do not allow to reconstruct the tensor. The practical implementation is straightforward, as the LS solution $\hat{\mathbf{d}}_{LS}$ is given by:

$$\hat{\mathbf{d}}_{LS} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}$$

The linear Least Squares method is further referred to as LS.

2.2.2 Weighted least squares

However, this method doesn't take into account the characteristics of noise. Indeed, as depicted in figure 2.1, due to the non-linearity of the equation of Stejskal and Tanner, the same uncertainty on the raw measurements η isn't propagated linearly on the corresponding estimated diffusivity: the induced uncertainty ϵ on the diffusivity is higher for low measurements. The lower the measured signal, the less accurate.

Basser *et al.* [6] take into account the effects of the log-transformation, weighting the square errors in the cost function by the inverse of the covariance matrix of the log-measurements. They carried out an error propagation, as explained in detail by Bevington and Robinso in [9]. The conclusion is that if the variance of hte measurements is σ_k^2, then the variance of the log-measurement y_k is $\sigma_k^2/b^2 s_k^2$. Basser *et al.* propose a weighted version of least squares, where the contribution of the data record are weighted by the inverse of the covariance matrix. As the measurements are uncorrelated, the covariance matrix $\mathbf{\Sigma}$ is diagonal and the estimate of the parameters of the tensor $\hat{\mathbf{d}}_{WLS}$ is now defined as:

$$\begin{aligned}\hat{\mathbf{d}}_{WLS} &= \arg\min_{\mathbf{d}} (\mathbf{H}\mathbf{d} - \mathbf{y})^T \mathbf{\Sigma}^{-1} (\mathbf{H}\mathbf{d} - \mathbf{y}) \\ &= (\mathbf{H}^T\mathbf{\Sigma}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{\Sigma}^{-1}\mathbf{y}\end{aligned}$$

In the remaining of this paper, WLS will stand for this estimate.

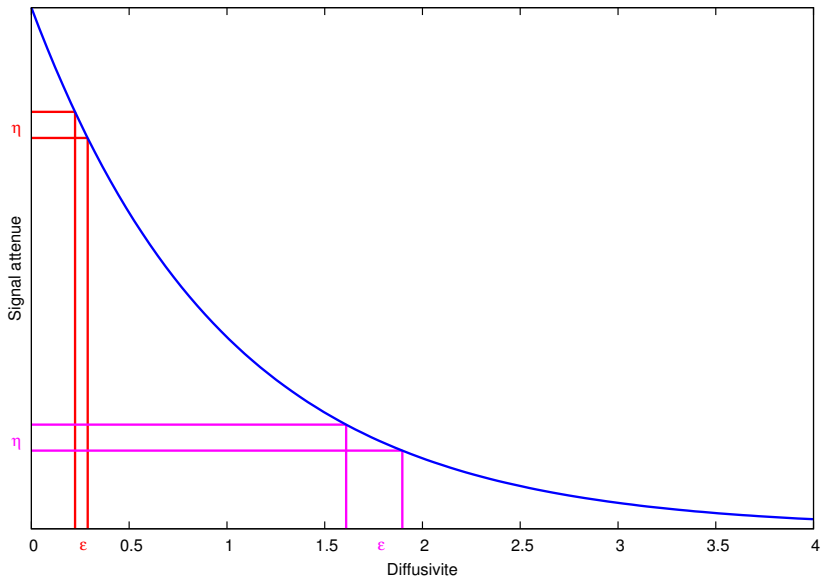


Figure 2.1: Non linearity of the Stejskal and Tanner equation

2.2.3 Iterated Least Squares

As an alternative to error propagation, we calculate the first order

$$S_k = S_0 e^{-b\mathbf{h_k}^T\mathbf{d}} + \eta_k$$

which yields, taking the logarithm:

$$-\frac{1}{b}\ln\left(\frac{S_k}{S_0}\right) \simeq \mathbf{h_k}^T\ \mathbf{d} + \frac{\eta}{bS_0 e^{-b\mathbf{h_k}^T\mathbf{d}}}$$

Then when this first order approximation holds, we can consider the noise ϵ_k on the log-measurements to be additive, and even Gaussian. However the variance is modified:

$$\text{if } \eta_k \sim \mathcal{N}(0, \sigma_k^2), \text{then } \epsilon \sim \mathcal{N}(0, \frac{\sigma_k^2}{(bS_0 e^{-b\mathbf{h_k}^T\mathbf{d}})^2} \tag{2.1}$$

As a refinement to WLS, we also introduce a weighted version of LS, but where the weights are calculated as the inverse of the expected variance of the log-measurements. Relying on the derivation in the previous paragraph, equation 2.1 suggests that the weights depend themselves on the tensor to be estimated. So we propose an iterative version of least squares estimation, which we call ILS, where the weights are calculated using the previous estimate of $\mathbf{d}$.

Actually, the weights at step $i+1$ are given by

$$\frac{1}{\sigma^2(\epsilon_k)^{(i+1)}} = \frac{(bS_0 e^{-b\mathbf{h_k}^T\hat{\mathbf{d}}^{(i)}})^2}{\sigma_k^2}$$

where $\hat{\mathbf{d}}^{(i)}$ is the tensor estimate at step i. The reweighting algorithm is presented in figure 2.2. We start with the WLS estimate as an initial guess: $\hat{\mathbf{d}}^{(0)} = \hat{\mathbf{d}}_{LS}$. With this approach, the weighting function is driven by the current estimate of the model, and then benefits from information given by all the measurements. As a consequence, we can expect the ILS to be more robust to noise than WLS, which is confirmed by the results shown in the next chapter.

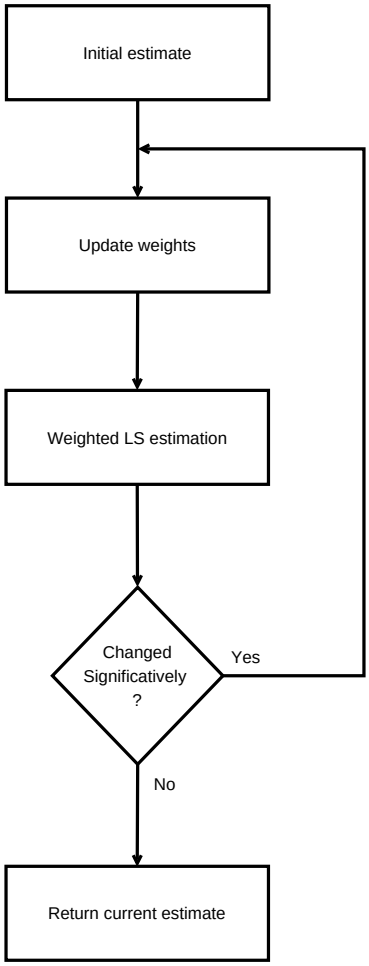


Figure 2.2: Reweighting algorithm

2.3 Non-linear tensor estimation

Although above mentioned methods are very computationally attractive, they assume moderate noise level, and are inadequate for very low SNR. In such situations, we should explicitly take into account more precisely the Rician nature of noise.

2.3.1 Non-linear least squares

We saw in the first section that the measured signal could be modelled by a complex signal, corrupted by Gaussian noise. Moreover, Wang *et al.* [10] show that complex DWI also obey the Stejskal and Tanner equation. They estimate $\mathbf{D}$ as the minimizer of an energy function with a regularization and an attachment term. The latter is a least of squares on the measured signal, which is actually the maximum likelihood estimate under the assumption of Gaussian additive noise:

$$\hat{\mathbf{D}} = \hat{\mathbf{L}}\hat{\mathbf{L}}^T \quad \text{where } \hat{\mathbf{L}} = \arg\min_{\mathbf{L}} \sum_{k=1}^{N} (R_k - R_0 e^{-b g_k^T \mathbf{L}\mathbf{L}^T g_k})^2 + (I_k - I_0 e^{-b g_k^T \mathbf{L}\mathbf{L}^T g_k})^2$$

Note that they also propose a particular parameterisation of $\mathbf{D}$ using the Cholesky decomposition, in order to ensure the positive-definiteness of the estimate. This approach is statistically

somehow optimal, as the estimated tensor is a model predicting (through Stejstal and Tanner equation) DWIs which are the minimum variance unbiased estimator (classical result for Gaussian distributed data). However, commonly one just can access the magnitude of the MR signal, hence the Rician PDF is to be considered.

2.3.2 Maximum likelihood estimation

In the section 1.3.1, we reminded the PDF of magnitude data:

$$p(S_k|A_k,\sigma) = \frac{S_k}{\sigma^2} e^{-\frac{S_k^2+A_k^2}{2\sigma^2}} I_0\left(\frac{A_k S_k}{\sigma^2}\right)$$

where S_k is the measured signal, and $A_k = S_0 e^{-b\mathbf{h_k}^T\mathbf{d}}$ is the expected signal magnitude. As this PDF involves complicated mathematical functions, it is not straightforward to find a minimal sufficient statistic. Following the method to find an adequate estimator described by Kay *et al.* [11], the natural way is to use the maximum likelihood estimator (MLE).

Principle of the MLE

The principle is easy and intuitive: given a data record, the PDF is seen as a function of the parameters, and is called the likelihood function. The MLE is the value of parameters maximizing this function, say it is the value of parameters that are the most likely, given the model and the data record.

At this point we can consider two different methods: we can either parameterize the likelihood function by the non-noisy value of the raw measurements A_k, or directly replace A_k by $S_0 e^{-b\mathbf{h_k}^T\mathbf{d}}$, and then the parameters are the coefficients of the diffusion tensor. Among the two methods, the former was proposed by Basu *et al.*, while the latter was experienced by Fillard *et al.*. We quickly review the two methods.

Denoising DWIs with a maximum-likelihood approach

In order to process images the less noisy as possible, it is common that various scans of the volume to be imaged are done, and then the mean value of these scans is returned by the scanner as the denoised image. For Gaussian distributed data, this technique is known to reduce the standard deviation by a factor of $\sqrt{M}$, where M is the number of repeated scans. The choice to use the mean value is driven by statistical optimality arguments, as the mean value of a vector of independant and identically-distributed random variables is the minimum variance unbiased estimator of the parameter of the Gaussian distribution.

The same processing can be done for images corrupted by Rician noise, in [2] the vector $\mathbf{A}$ of N DWIs used for tensor reconstruction is itself reconstructed from the noisy measurements $S_{k,l}$ as the minimizer of the energy:

$$U(\mathbf{A}) = \int_\Omega \sum_{k=1}^{N}\sum_{l=1}^{M} \frac{A_k^2}{2\sigma^2} - \log I_0\left(\frac{S_{k,l}A_k}{\sigma^2}\right) + \lambda c(||\nabla \mathbf{A}||^2)\ dx\ dy$$

If we consider that the various measurements are independant, then the join PDF is simply the product of each PDF. Forgetting about the regularization term, we have the MLE of $\mathbf{A}$, from which the tensor can be estimated using a linear regression.

Direct estimation of the tensor

An alternative method is to rewrite the likelihood function, so that it depends on $\mathbf{D}$. Fillard *et al.* in [3] implemented this technique, together with a regularization term. In order to impose the positive-definite constraint, they use a log-euclidian metric, which consists in estimating the log of the diffusion tensor. Writing $\mathbf{D} = \exp(\mathbf{L})$, the energy they minimize is:

$$U(\mathbf{L}) = \sum_{k=1}^{N} \frac{\tilde{S}_k(\mathbf{L})^2}{2\sigma^2} - \log I_0 \left(\frac{\tilde{S}_k(\mathbf{L}) S_k}{\sigma^2} \right) \quad \text{où } \tilde{S}_k(\mathbf{L}) = S_0 e^{-b g_k^T \exp(\mathbf{L}) g_k}$$

As an extension of the work of Lenglet *et al.* [12], we propose to minimize this function with a Riemannian gradient descent, also in order to ensure the positiveness of the estimated tensor. We hereafter derive the expression of the Riemannian gradient, and remind the numerical scheme used for minimization.

A Riemannian approach

In the expression of the log-likelihood appear terms not depending on the parameters to be estimated, so after simplificating, the energy function related to the likelihood is $U_{ML}(\mathbf{D}) = \sum_{k=1}^{N} f_k(\mathbf{D})$, where:

$$f_k(\mathbf{D}) \quad = \quad \frac{S_0^2}{2\sigma^2} e^{-2b g_k^T \mathbf{D} g_k} + \log I_0 \left(\frac{S_k S_0}{\sigma^2} e^{-b g_k^T \mathbf{D} g_k} \right)$$

We need to find the expression of $\nabla_{\mathbf{D}} f_k$, namely the Riemannian gradient of f_k. If we consider a geodesic curve $\mathbf{S}(t) = \mathbf{D}^{1/2} \exp(t\mathbf{A}) \mathbf{D}^{1/2}$ starting from $\mathbf{D}$, with the tangent direction $\mathbf{\Delta} = \dot{\mathbf{S}}(0) = \mathbf{D}^{1/2} A \mathbf{D}^{1/2}$, then the gradient of f_k at $\mathbf{D}$ should verify :

$$\frac{\mathrm{d}}{\mathrm{d}t} f_k\left(\mathbf{S}(t)\right)\Big|_{t=0} = \langle \mathbf{\Delta}, \nabla_{\mathbf{D}} f_k \rangle_{\mathbf{D}} \tag{2.2}$$

where the inner product on the tangent space at $\mathbf{D}$ is defined by $\langle \mathbf{A}, \mathbf{B} \rangle_{\mathbf{D}} = \mathrm{Tr}(\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-1}\mathbf{B})$

So we first calculate the derivative of $f_k(\mathbf{S}(t))$, with respect to t:

$$\frac{\mathrm{d}}{\mathrm{d}t} f_k\left(\mathbf{S}(t)\right) = \left(\frac{S_k S_0 b}{\sigma^2} \; e^{-b\mathrm{Tr}(\mathbf{S}(t)\mathbf{G_i})} \; \frac{I_0'}{I_0} \left(\frac{S_k S_0}{\sigma^2} \; e^{-b\mathrm{Tr}(\mathbf{S}(t)\mathbf{G_k})} \right) - \frac{S_0^2 b}{\sigma^2} \; e^{-2b\mathrm{Tr}(\mathbf{S}(t)\mathbf{G_k})} \right) \mathrm{Tr}(\dot{\mathbf{S}}(t)\mathbf{G_k}) \tag{2.3}$$

Writing $\mathbf{G_k} = g_k g_k^T$ and taking the value for $t = 0$, we have :

$$\frac{\mathrm{d}}{\mathrm{d}t} f_k\left(\mathbf{S}(t)\right)\Big|_{t=0} = \left(\frac{S_k S_0 b}{\sigma^2} \; e^{-b\mathrm{Tr}(\mathbf{D}\mathbf{G_i})} \; \frac{I_0'}{I_0} \left(\frac{S_k S_0}{\sigma^2} \; e^{-b\mathrm{Tr}(\mathbf{D}\mathbf{G_k})} \right) - \frac{S_0^2 b}{\sigma^2} \; e^{-2b\mathrm{Tr}(\mathbf{D}\mathbf{G_k})} \right) \mathrm{Tr}(\mathbf{\Delta}\mathbf{G_k}) \tag{2.4}$$

Finally, the Riemannian gradient is :

$$\nabla_{\mathbf{D}} f_i = \left(\frac{S_k S_0 b}{\sigma^2} \; e^{-b\mathrm{Tr}(\mathbf{D}\mathbf{G_i})} \; \frac{I_0'}{I_0} \left(\frac{S_k S_0}{\sigma^2} \; e^{-b\mathrm{Tr}(\mathbf{D}\mathbf{G_k})} \right) - \frac{S_0^2 b}{\sigma^2} \; e^{-2b\mathrm{Tr}(\mathbf{D}\mathbf{G_k})} \right) \mathbf{D}\mathbf{G_k}\mathbf{D} \tag{2.5}$$

Lenglet *et al.*propose a step-forward operator to minimize a function, while staying within the set of positive symmetric matrices. So we take as an initial guess the result of the LS estimation (which we may possibly reproject into the set of positive matrices), and then we apply the scheme:

$$\mathbf{D}_{l+1} = \mathbf{D}_l^{1/2} \exp(-dt \mathbf{D}_l^{-1/2} \nabla U_{ML} \mathbf{D}_l^{-1/2}) \mathbf{D}_l^{1/2}$$

Connexion to the linear approaches

As we showed that for sufficiently high SNR, Gaussian and Rician noise were similar, it may be interesting to verify that the techniques in use for Rician noise become similar to linear techniques at high SNR.

If we write the likelihood function, replacing the bessel function $I_0\left(\frac{A_k S_k}{\sigma^2}\right)$ by its asymptotic equivalent (for $\frac{A_k S_k}{\sigma^2} \gg 1$) $\exp(\frac{A_k S_k}{\sigma^2})/\sqrt{2\pi \frac{A_k S_k}{\sigma^2}}$, we have:

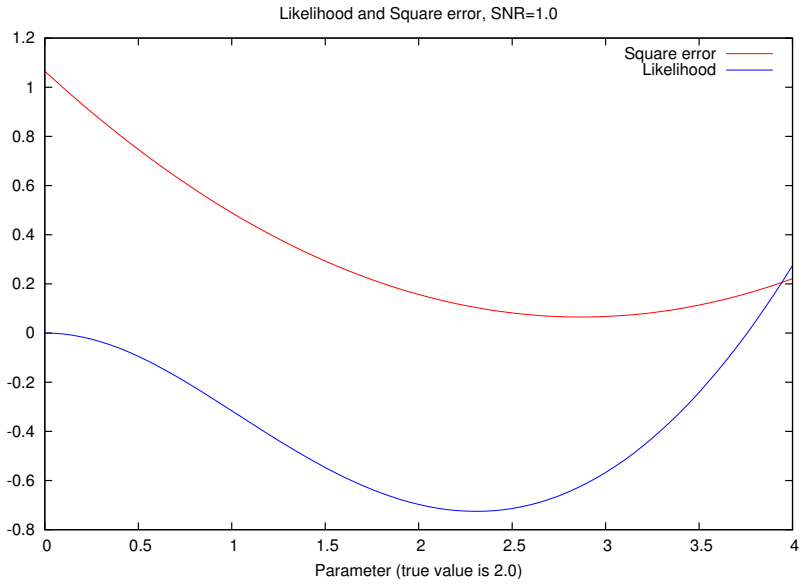
$$p(S_k|A_k,\sigma) = \frac{S_k}{\sigma^2}\exp(-\frac{S_k^2 + A_k^2}{2\sigma^2})\frac{\exp(\frac{A_k S_k}{\sigma^2})}{\sqrt{2\pi \frac{A_k S_k}{\sigma^2}}}$$

and after simplifying, we find again the Gaussian pdf:

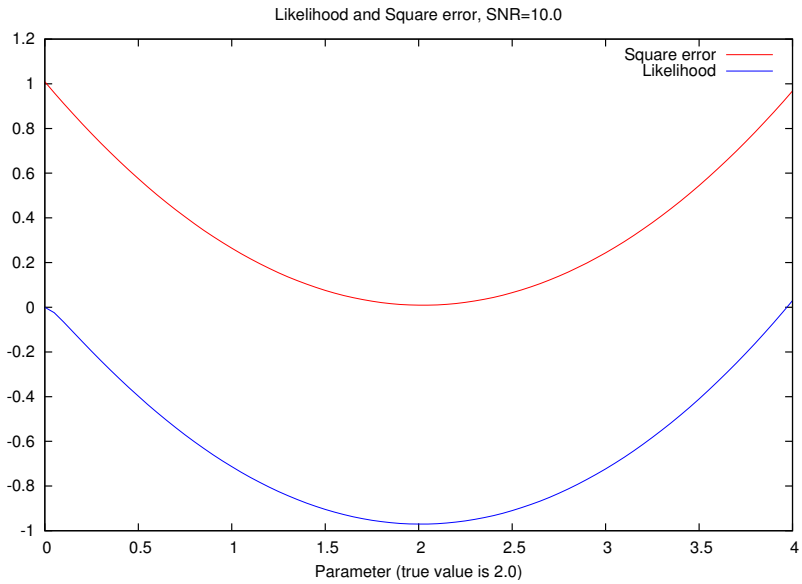
$$p(S_k|A_k,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(S_k - A_k)^2}{\sigma^2})$$

Maximizing the likelihood is then the same for Rician and for Gaussian distributed data. The plots in figure 2.3 correspond to the energy to be minimized, when estimating the A parameter of the Rician distribution, from a series of 10 measurements randomly generated following the Rice distribution. When the SNR equals 10, the energy corresponding to LS estimation is just a scaled version of the maximum likelihood-related energy, and hence the estimates are equal.

The techniques presented in this chapter (namely LS, WLS, ILS and MLE) were implemented. The next section first presents the way our synthetic data were generated, and then presents some results on it.



(a) Low SNR



(b) High SNR

Figure 2.3: Energy function corresponding to LS and maximum-likelihood.

Chapter 3

Comparison and statistics on reconstruction techniques

The aim of this section is to compare the classical estimation techniques (classified into two categories, linear and non linear estimation), and to evaluate the contribution of non-linear methods. As the accuracy of each technique is strongly linked with the level of noise, the objective was also to carry out a quantitative study on SNR.

3.1 Methods

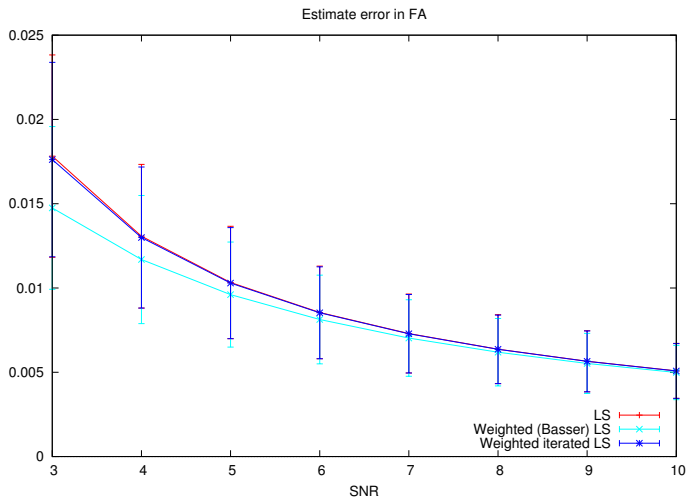
We would like to get statistics on the estimated tensor, for the various techniques previously described. Here is the way we generated our raw data: we started with a synthetic tensor field, uniform, which is considered as a the ground truth. Practically it is a set of 8000 tensors (a volume of $20 \times 20 \times 20$). Given a set of 81 directions (generated with the tesselation method of the sphere, as used by Tuch in [13]), we generated the corresponding DWIs, following the Stejskal and Tanner equation, with a b-value equal to 3000. Then some Rician noise was added on the DWIs, as the magnitude of complex data corrupted by Gaussian additive noise.

The test were driven on two different kind of tensors, with eigenvalues similar to what can be measured in real tissues; the eigenvalues of the anisotropic tensor were (in mm^2/s) $(1.3\ 10^{-3}, 2.3\ 10^{-4}, 2.3\ 10^{-4})$, and for the isotropic tensor $(2.0\ 10^{-3}, 2.0\ 10^{-3}, 2.0\ 10^{-3})$. The values of SNR range from 2.0 to 10.0; such values may look unlikely low, however in condition of very high b-value, it is common to record such noisy data. The SNR is defined as the ratio between the mean signal on the whole set of DWIs:

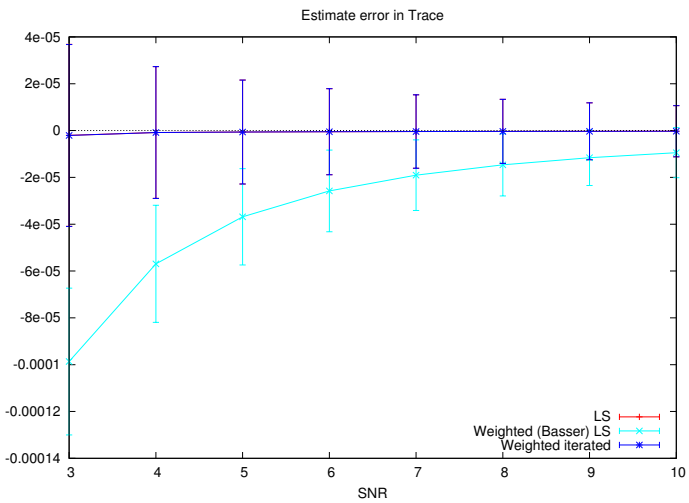
$$\mathrm{SNR} = \frac{1}{N|\Omega|} \frac{\sum_{k \in \{1 \ldots N\}, s \in \Omega} S_k(s)}{\sigma}$$

and we assume the σ parameter to be uniform on the volume.

The evaluation of performance was done through statistics on the scalar indices extracted from the tensor: FA and trace, as well as on the principal direction of the estimated tensor (the latter was only calculated for the anisotropic tensor, as the isotropic tensor has no preferred direction). The plots correspond to the mean and standard deviation of error with respect to the ground truth. Negative values suggest that the corresponding value (FA or trace) is underestimated, while positive values correspond to a positive bias.



(a) FA



(b) Trace

Figure 3.1: Error statistics: isotropic tensor, comparison of LS, WLS and ILS

3.2 LS and Weighted LS

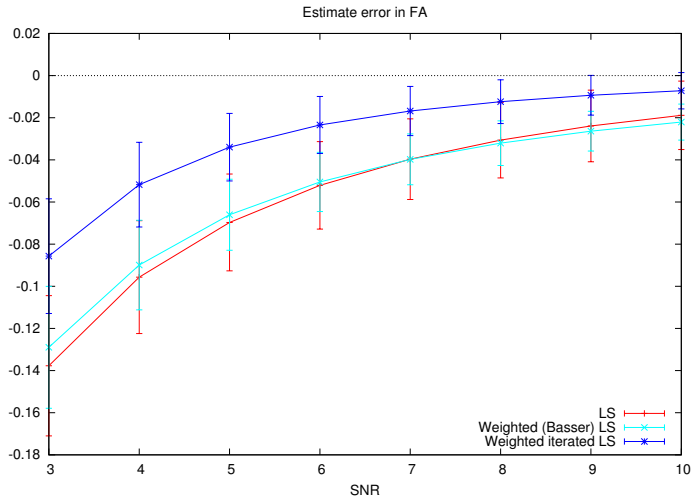
In this section we show the results of linear estimation techniques: we compare linear LS (presented in section 2.2.1), WLS (section 2.2.2), and ILS (section 2.2.3). The results on isotropic tensor in figure 3.1 show that ILS performs as well as LS, which is intuitive: for isotropic tensor, the contribution of each measurement is expected to be similar, so our weighting is approximately equal for all measurements. Conversely, as WLS method uses directly the measured signal for the ponderation, it is more sensitive to noise, and then show results even poorer than without weighting for this particular case.

For an anisotropic tensor (see figure 3.2 for results), WLS performs better than classical LS, as it tends to reduce the influence of low (and hence noisy and biased) measurements. Our technique offers better results, as it uses a loop from the estimated tensor to robustly calculate the approriate weights.

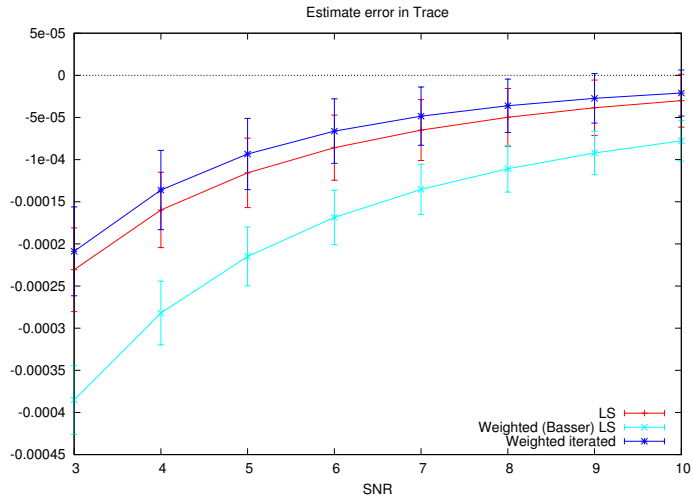
3.3 Maximum-likelihood estimation

The MLE was implemented, with the Riemannian gradient approach for energy minimization. As well as for the linear techniques, we plot the error in FA and Trace, and compare to the ILS algorithm (which performs the best among linear techniques on these error metrics). The plots are presented in figure 3.3. Note that we didn't include the MLE error plot on the previous figures, as we wanted to keep them as clear as possible, and also the interesting range of SNR is different.

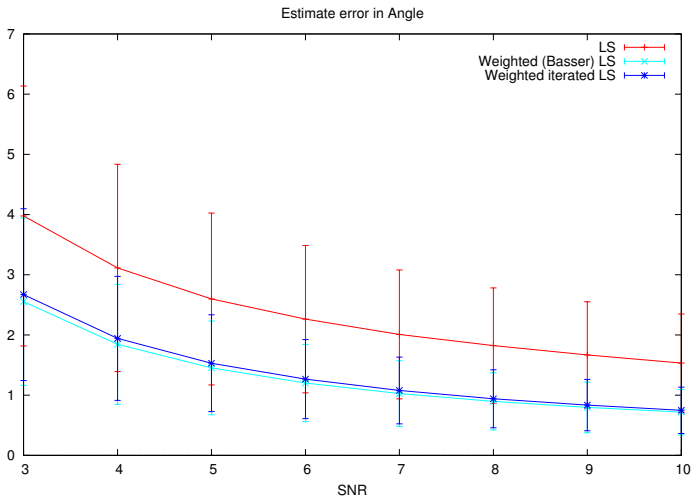
The quantitative results confirm the qualitative observations we did previously. Both the FA and the Trace are underestimated by linear techniques (even our weighted LS criterion doesn't compensate the bias), although the MLE can retrieve them accurately. This study also demonstrates how linear and non linear techniques become equivalent for a sufficiently high SNR.



(a) FA

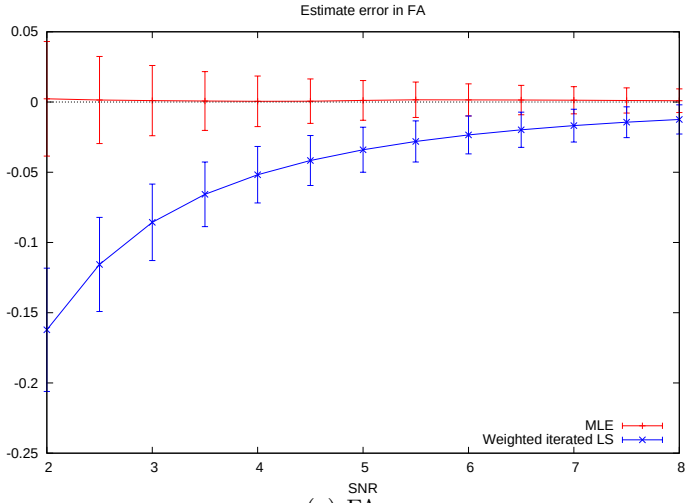


(b) Trace

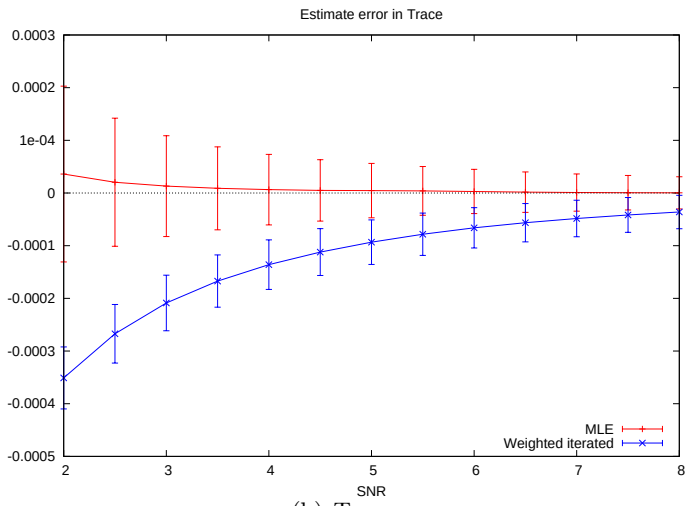


(c) Angle

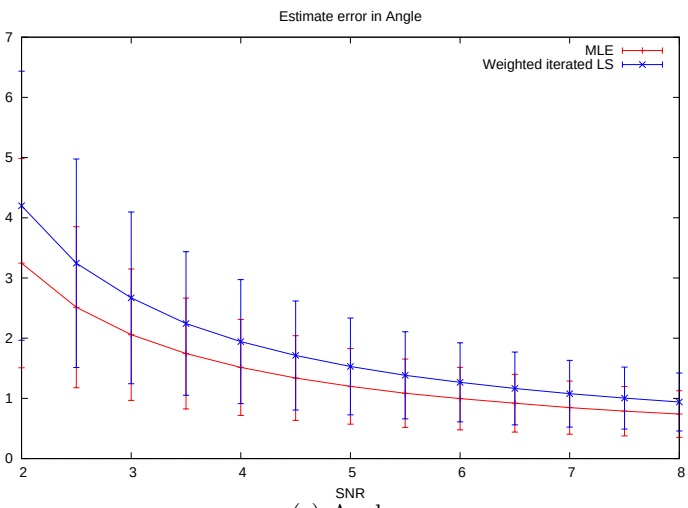
Figure 3.2: Error statistics: anisotropic tensor, comparison of LS, WLS and ILS



(a) FA



(b) Trace



(c) Angle

Figure 3.3: Error statistics: anisotropic tensor, comparison of ILS and MLE.

Chapter 4

Conclusion and discussion

We have tested various estimation methods for diffusion tensor reconstruction, classified into two classes: either linear or non-linear. We showed that non-linear techniques like the MLE can significantly improve the results, with the drawback of a very high computational cost. So far, and overall for the scope of clinical applications, the processing time tends to be the decisive factor for choosing a method instead of an other.

Hence the result on WLS and ILS are of particular interest, as the computational cost remains very close to classical LS. Moreover, this technique can easily be adapted to reconstruction of high angular model reconstruction, such as an Orientation Distribution Function (ODF) reconstruction. For these models, there are too many parameters (15 when a spherical harmonics basis of order 4 is used) to consider minimizing a non-linear function in a reasonable time. So far, the reconstruction method of the ODF proposed by Descoteaux *et al.* [14] is a linear fit on the measurements, with no weighting at all. We may benefit from the results we have on tensor estimation to propose a weighted version of ODF reconstruction from raw measurements. If the extrapolation is accurate, as the reconstructed tensor are sharper we can think of this method for reconstructing sharper ODFs too.

Bibliography

[1] J Sijbers. *Signal and noise estimation from magnetic resonance images.* PhD thesis, Universiteit Antwerpen, October 1998.

[2] S Basu, P T Fletcher, and R T Whitaker. Rician noise removal in diffusion tensor mri. In *MICCAI (1)*, pages 117–125, 2006.

[3] P Fillard, V Arsigny, X Pennec, and N Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with Log-Euclidean metrics. *IEEE Transactions on Medical Imaging*, 2007. In Press.

[4] E O Stejskal and J E Tanner. Spin diffusion measurements: spin echoes in the presence of time-dependant field gradient. *J. Chem. Phys.*, 42:288–292, 1965.

[5] D Le Bihan, E Breton, D Lallemand, P Grenier, E Cabanis, and M Laval-Jeantet. Mr imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. *Radiology*, pages 401–407, 1986.

[6] P J Basser, J Mattiello, and D LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophys J*, 66(1):259–267, Jan 1994.

[7] K M Hasan, D L Parker, and A L Alexander. Comparison of gradient encoding schemes for diffusion-tensor MRI. *J Magn Reson Imaging*, 13(5):769–780, May 2001. Comparative Study.

[8] R Rathore. Necessary and sufficient conditions for the admissibility of dti gradient vectors. In *Proceedings of the Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM 2007)*, Berlin (Germany), May 2007.

[9] P R Bevington. *Data reduction and error analysis for the physical sciences.* McGraw-Hill, New-York, 1969.

[10] Z Wang, B C Vemuri, Y Chen, and T H Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex dwi. 23(8):930–939, August 2004.

[11] S M Kay. *Fundamentals of statistical signal processing: estimation theory.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

[12] C Lenglet. *Geometric and Variational Methods for Diffusion Tensor MRI Processing.* PhD thesis, Université de Nice-Sophia Antipolis, December 2006.

[13] D S Tuch. *Diffusion MRI of Complex Tissue Structure.* PhD thesis, massachusetts institute of technology, 2002.

[14] M Descoteaux, E Angelino, S Fitzgibbons, and R Deriche. Regularized, fast and robust analytical q-ball imaging. *Magn. Res. in Med. (to appear)*, 2007.

Physique de la formation des images en IRM de diffusion
Rapport bibliographique

Emmanuel Caruyer, Master de Recherche en Informatique.

1er février 2007

Ce document constitue un pré-rapport pour mon stage de Master, effectué au sein de l'équipe Odyssée à l'INRIA Sophia-Antipolis et encadré par Rachid Deriche. Je tiens à remercier Rachid Deriche et Maxime Descoteaux pour avoir accepté de m'accueillir dans leur équipe, pour le temps qu'ils ont passé à m'introduire le sujet et pour leur aide précieuse concernant ce rapport.

Introduction

La découverte du phénomène de résonnance magnétique nucléaire a permis de développer un outil révolutionnaire dans le domaine de l'imagerie médicale. L'Imagerie par Résonnance Magnétique (IRM) en médecine a significativement contribué à améliorer la compréhension du corps, et constitue par ailleurs un outil unique pour le diagnostic de certaines pathologies. Le développement conjoint des technologies informatiques et de modèles mathématiques adaptés a permis de mettre en oeuvre des algorithmes de traitement du signal *ad hoc* pour exploiter au mieux les données mesurées.

Au milieu des années 1980, Le Bihan *et al.* ont montré qu'il était possible de rendre l'IRM sensible au mouvement des protons contenus dans les molécules d'eau, et ainsi de remonter à la structure sous-jacente des tissus. Quelques années plus tard, Basser *et al.* proposent un formalisme tensoriel qui permet de décrire l'anisotropie de la diffusion des molécules d'eau. Depuis, plusieurs travaux ont été menés sur l'estimation, la régularisation et l'utilisation du champ de tenseurs de diffusion ; et plus récemment, les chercheurs ont remis en cause le modèle tensoriel pour proposer des modèles à haute résolution angulaire, d'un intérêt particulier pour l'observation des fibres de la matière blanche.

Notre étude bibliographique s'est concentrée sur ces derniers travaux, ainsi que sur la caractérisation du bruit en IRM. Nous présentons dans un premier temps un panorama des modèles de diffusion moléculaire et leur utilisation en imagerie médicale. Dans la seconde partie, quelques rappels sur la physique de l'IRM sont présentés, afin d'introduire le principe de l'IRM de diffusion et de décrire les sources de bruit dans le signal mesuré. La troisième et dernière partie est consacrée aux algorithmes d'estimation et de régularisation des paramètres des différents modèles utilisés.

1 Un modèle pour capturer la diffusion moléculaire

La diffusion caractérise le mouvement spontané des molécules d'eau sous l'effet de l'agitation thermique. Ce mouvement est contraint par la structure du milieu sous-jacent ; dès lors, en mesurant la diffusion à l'échelle macroscopique (la taille du voxel en imagerie médicale est typiquement de 1mm $\times$ 1mm $\times$ 2mm), on peut remonter à des informations sur la structure fine des tissus. Ceci constitue l'enjeu majeur de l'étude du phénomène de diffusion en imagerie médicale, et ce particulièrement pour l'observation de la matière blanche du cerveau, constituée d'un réseau de fibres neuronales assurant la connectivité cérébrale. La diffusion est un phénomène complexe, plusieurs modèles pour le décrire ont été proposés, nous les présentons dans l'ordre chronologique des publications qui les décrivent.

1.1 Le tenseur de diffusion

L'imagerie tensorielle de diffusion [1] propose d'estimer un tenseur en chaque voxel. La diffusion en un voxel est caractérisée par la densité de probabilité p du vecteur $\mathbf{x}$, déplacement d'une molécule pendant une certaine durée t. Le tenseur de diffusion $\mathbf{D}$ est en fait la matrice de covariance de cette variable aléatoire $\mathbf{x}$, et en admettant que $\mathbf{x}$ suit une loi gaussienne, on peut décrire la diffusion par ce seul tenseur. Le signal est sensible à la diffusion dans une direction $\mathbf{g}_k$ en appliquant un *gradient de diffusion* (voir partie 2 pour plus de détails) dans cette direction. L'atténuation observée par rapport au signal mesuré sans gradient de diffusion suit l'équation de Stejskal-Tanner :

$$S(k) = S(0)\exp(-b\mathbf{g}_k^T\mathbf{D}\mathbf{g}_k) \tag{1}$$

où $S(0)$ est le signal sans gradient de diffusion, et $\mathbf{D}$ est le tenseur de diffusion.

Le tenseur peut être représenté par un ellipsoïde, et des composantes spectrales (vecteurs propres et valeurs propres λ_i) on peut extraire des informations synthétiques sur la diffusion. En particulier, la trace du tenseur $Tr(\mathbf{D})$ correspond au déplacement quadratique moyen des molécules d'eau ; alors que l'anisotropie fractionnelle ν :

$$\nu \propto \frac{\sum_{i=1}^{3}(\lambda_i - \overline{\lambda_i})^2}{\sum_{i=1}^{3}\lambda_i^2}$$

quantifie la « structuration » du milieu. En imagerie cérébrale, les plus fortes anisotropies sont relevées dans la matière blanche, là où les fibres nerveuses sont regroupées en faisceaux et où les molécules d'eau ont un mouvement cohérent. Dans ces zones fortement anisotropiques, la diffusion est privilégiée le long des fibres et le vecteur propre associé à la valeur propre principale du tenseur $\mathbf{D}$ donne la direction du faisceau de fibres sous-jacent. Sur ces observations, des algorithmes de tractographie des faisceau de fibres de la matière blanche ont été développés, afin de reconstruire la connectivité cérébrale.

Ainsi le modèle tensoriel permet une représentation compacte du phénomène de diffusion moléculaire, et sa grande popularité du point de vue clinique réside dans l'intérêt des cliniciens pour les cartes d'anisotropie et de diffusion moyenne. En revanche, l'utilisation du modèle tensoriel en tractographie soulève des doutes ; un effet de volume partiel est observé lors du croisement de faisceaux de fibres au sein d'un même voxel, et le vecteur propre associé à la valeur propre principale ne caractérise plus fiablement la direction des fibres.

1.2 Modèle de diffusion multimodal

La diffusion moléculaire au sein d'un même voxel s'effectue dans des milieux différents (typiquement en milieux intra-cellulaire et extra-cellulaire), ou aux caractéristiques différentes ; il a par exemple été mis en évidence par Clark et Le Bihan [2] que la diffusion n'est pas aussi rapide dans chacun de ces milieu (typiquement plus lente à l'intérieur des cellules qu'à l'extérieur). En supposant qu'il y ait peu d'échanges entre ces milieux, on peut alors décrire la diffusion globale comme une combinaison de la diffusion dans chacun de ces sous-milieux.

Ce modèle constitue un raffinement du modèle tensoriel, mais cependant il est difficile de mettre en oeuvre un algorithme de reconstruction s'appuyant sur ce modèle. Des problèmes de stabilité peuvent apparaître si l'on cherche à estimer deux tenseurs pour un site monomodal ; Alexander *et al.* [3] ont étudié une solution adaptative, basée sur un test d'hypothèse de pertinence d'un modèle à deux milieux contre un modèle tensoriel classique. Cette technique permet donc d'éviter les problèmes de stabilité, mais est difficilement généralisable à plus de deux tenseurs.

1.3 Estimation du profil du coefficient apparent de diffusion

Pour s'affranchir des limitations du modèle tensoriel, on ne suppose plus que la diffusion est gaussienne, on s'intéresse à une gamme plus large de *pdf* pour $\mathbf{x}$. Le profil du coefficient apparent de diffusion est une fonction sphérique ADC, qui caractérise la diffusion apparente pour chaque direction $\mathbf{g} = \mathbf{g}(\theta, \phi)$. Le coefficient apparent de diffusion peut être relié directement aux mesures effectuées ; le produit tensoriel $\mathbf{g}_k^T \mathbf{D} \mathbf{g}_k$ dans l'équation 1 est remplacé par $ADC(\mathbf{g}_k)$. Dans la littérature, deux espaces de projection ont été utilisés pour décrire l'ADC, à savoir les fonctions harmoniques sphériques et les tenseurs d'ordre supérieur.

1.3.1 Fonctions harmoniques sphériques

Les harmoniques sphériques forment une base orthonormale (pour le produit scalaire canonique $\langle f_1, f_2 \rangle = \int_S f_1^*(\mathbf{g}) f_2(\mathbf{g}) \mathrm{d}\mathbf{g}$) des fonctions complexes sur la sphère unité S. Ces fonctions, notées $Y_l^m, l \in \mathcal{N}, m \in \{-l, \ldots, l\}$ sont solutions de l'équation $\Delta_b F = -l(l+1)F$, où Δ_b est l'opérateur de Laplace-Beltrami, et peuvent être explicitées :

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) \exp(\imath m \phi)$$

où l représente l'ordre, m représente la phase, et P_l^m est un polynôme de Legendre. Descoteaux *et al.* [4] utilisent une version modifiée de cette base, pour décrire uniquement des fonctions sphériques réelles, positives et symmétriques (invariant par passage à l'antipode), contraintes imposées par la physique du phénomène de diffusion moléculaire. L'intérêt de cette base de projection est d'une part, comme montré dans la partie sur la régularisation, de s'adapter très bien au calcul d'un critère de régularité, dans l'optique de proposer une méthode de débruitage ; et d'autre part de proposer une mesure d'anisotropie, à l'instar de ce qui a été fait pour le modèle tensoriel.

1.3.2 Tenseur d'ordre supérieur

Cette proposition est une généralisation du modèle tensoriel classique, et ne présuppose plus une diffusion gaussienne. Le tenseur $\mathbf{D}$ d'ordre 2 dans l'équation 1 est remplacé par un tenseur d'ordre supérieur. Ce modèle est équivalent théoriquement et algorithmiquement au modèle sphérique harmonique. Descoteaux *et al.* [4] dérivent sous forme linéaire une transformation des

coefficients du modèle des fonctions harmoniques vers le modèle des tenseur d'ordre supérieur, ce qui permet de comparer les algorithmes de reconstruction et de régularisation proposés pour chacun, ainsi que les mesures d'anisotropie.

L'intérêt de telles modélisations permet de régulariser l'ADC reconstruit, et de caractériser plus finement l'anisotropie. Après régularisation, les zones correspondant à un profil de diffusion complexe peuvent être détectées facilement (elles sont significativement mieux décrites avec un modèle d'ordre supérieur). Cependant le coefficient apparent de diffusion reste très lié au signal mesuré ; il est probable qu'un modèle plus abstrait, plus intrinsèquement lié au phénomène de diffusion et à la *pdf* soit plus descriptif, et permette une meilleure détection des faisceaux de fibres par exemple.

1.4 Orientation density function

L'*odf* Ψ (pour *orientation density function*) est la projection radiale de la *pdf* : pour tout vecteur unitaire $\hat{\mathbf{q}}$, on a

$$\Psi(\hat{\mathbf{q}}) = \int_0^\infty p(\alpha\hat{\mathbf{q}})\mathrm{d}\alpha$$

Cette transformation de la *pdf* permet de décrire la structure angulaire de la diffusion, et en pratique les maxima de cette fonction Ψ correspondent aux directions privilégiées de diffusion (le long des fibres neuronales), ce qui permet d'accroître la sensibilité pour la détection de plusieurs fibres en un même voxel. Tuch *et al.* ont par ailleurs montré qu'il était possible d'estimer cette fonction Ψ à partir d'une séquence de mesures effectuées en haute résolution angulaire (imagerie $\mathbf{q}$-ball), en utilisant la transformée de Funk-Radon de l'ADC ; la transformée de Funk-Radon d'une fonction sphérique f en un point $\hat{\mathbf{q}}$ étant l'intégrale de f sur le grand cercle normal à $\hat{\mathbf{q}}$.

Plusieurs modèles se sont donc déclinés, pour répondre à un besoin intrinsèque (limitations du modèle tensoriel), et pour capturer l'intégralité de l'information mesurée. Nous avons présenté jusqu'ici les modèles pour décrire le phénomène de diffusion ; dans la partie qui suit nous nous intéressons à la façon dont l'IRM est rendue sensible au phénomène de diffusion.

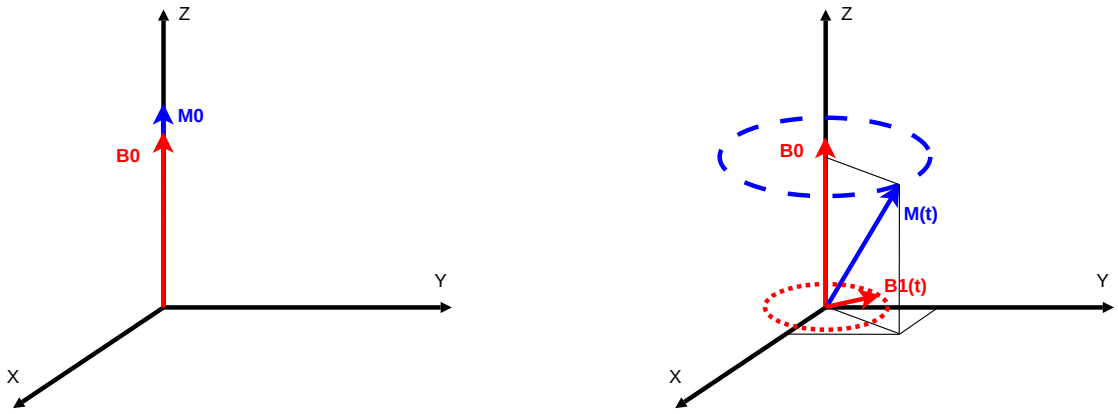
2 La physique du signal IRM de diffusion

Dans ce qui suit, nous commençons par rappeler les principes fondamentaux de l'IRM, puis expliquons comment le déplacement des molécules peut être mesuré en IRM.

2.1 Résonnance magnétique nucléaire

L'IRM est basée sur le phénomène de résonnance magnétique nucléaire : sous l'action d'un champ magnétique intense $\mathbf{B_0}$, un noyau magnétique tel que le proton a tendance à aligner son moment magnétique dans la direction du champ, et ainsi un moment magnétique macroscopique $\mathbf{M_0}$ apparaît. Le moment magnétique intrinsèque de chaque proton tourne en réalité autour de $\mathbf{B_0}$ à la fréquence de Larmor $\omega_0 = \gamma\|\mathbf{B_0}\|$, qui dépend du champ extérieur et du *coefficient gyromagnétique* γ.

En excitant les spin par un champ magnétique transverse $\mathbf{B_1}$ tournant à la fréquence de Larmor, les spins tournent alors en phase et le moment magnétique macroscopique résultant est modifié. Conventionellement, le champ constant $\mathbf{B_0}$ est suivant $\mathbf{z}$, ainsi le champ tournant est dans le plan $(\mathbf{x}, \mathbf{y})$ (*cf.* figure 1). Suivant la durée de l'impulsion d'excitation radiofréquence, on peut annuler la composante suivant $\mathbf{z}$ (impulsion 90°), ou même l'inverser (impulsion 180°). Le retour à l'équilibre thermique de Boltzmann se fait suivant les équations de Bloch (rappelées



(a) À l'équilibre (b) Pendant l'excitation

FIG. 1 – Réponse à l'excitation causée par le champ tournant.

dans [5]) :

$$\mathbf{M}_{x,y}(t) = \mathbf{M}_{x,y}(0) \exp\left(-\frac{t}{T_2}\right)$$

$$\mathbf{M}_z(t) = \mathbf{M}_z(0) \exp\left(-\frac{t}{T_1}\right) + \mathbf{M}_0 \left(1 - \exp\left(-\frac{t}{T_1}\right)\right)$$

T_1 est le temps de relaxation longitudinal, et T_2 le temps de relaxation transversal. Cette variation de moment magnétique créée un flux électromagnétique, que l'on peut mesurer à l'aide d'un courant induit dans une antenne (bobine placée autour de l'échantillon) ; le signal complexe mesuré s'appelle le *free induction decay* (*FID*).

2.2 Séquences d'acquisition

Le principe décrit au paragraphe précédent s'applique pour l'analyse d'échantillons homogènes ; l'utilisation du principe de résonnance magnétique en imagerie nécessite de pouvoir identifier les contributions au *FID* des différentes zones de l'espace.

2.2.1 Localisation du signal

On utilise le fait que la fréquence de Larmor dépend de l'intensité du champ ; en faisant varier spatialement l'intensité du champ (introduction d'un *gradient* de champ magnétique), on peut « signer » en fréquence le signal en fonction de l'endroit d'où il provient. Par le même principe, suivant le moment dans la séquence où le gradient est appliqué, on peut également modifier la phase de précession en fonction de la position (en appliquant un gradient pendant un court instant). Le choix de la fréquence d'émission du signal permet de sélectionner au préalable une coupe transversale (sachant qu'un gradient suivant $\mathbf{z}$ est appliqué), puis le codage en fréquence et en phase permet de localiser la provenance du signal au sein de la coupe par une transformée de Fourrier inverse du signal.

2.2.2 Images pondérées en diffusion

Pour mesurer un signal qui tienne compte du déplacement des molécules d'eau, on ajoute à la séquence d'écho de spin précédemment décrite deux impulsions de gradient selon la direction $\mathbf{g}_k$ (dit gradient de pondération en diffusion) identiques, juste avant et juste après l'impulsion

180° (impulsion d'inversion de phase). Ces deux impulsions n'ont aucun effet sur les particules fixes (le déphasage introduit par la première est compensé par la seconde). Par contre, le signal mesuré est atténué si les protons se déplacent dans la direction du gradient de diffusion pendant la durée Δ séparant les deux impulsions de gradient (*cf.* figure 2). En répétant cette séquence d'écho de spin pour différentes orientations du gradient de diffusion, on peut mesurer la diffusion dans différentes directions.

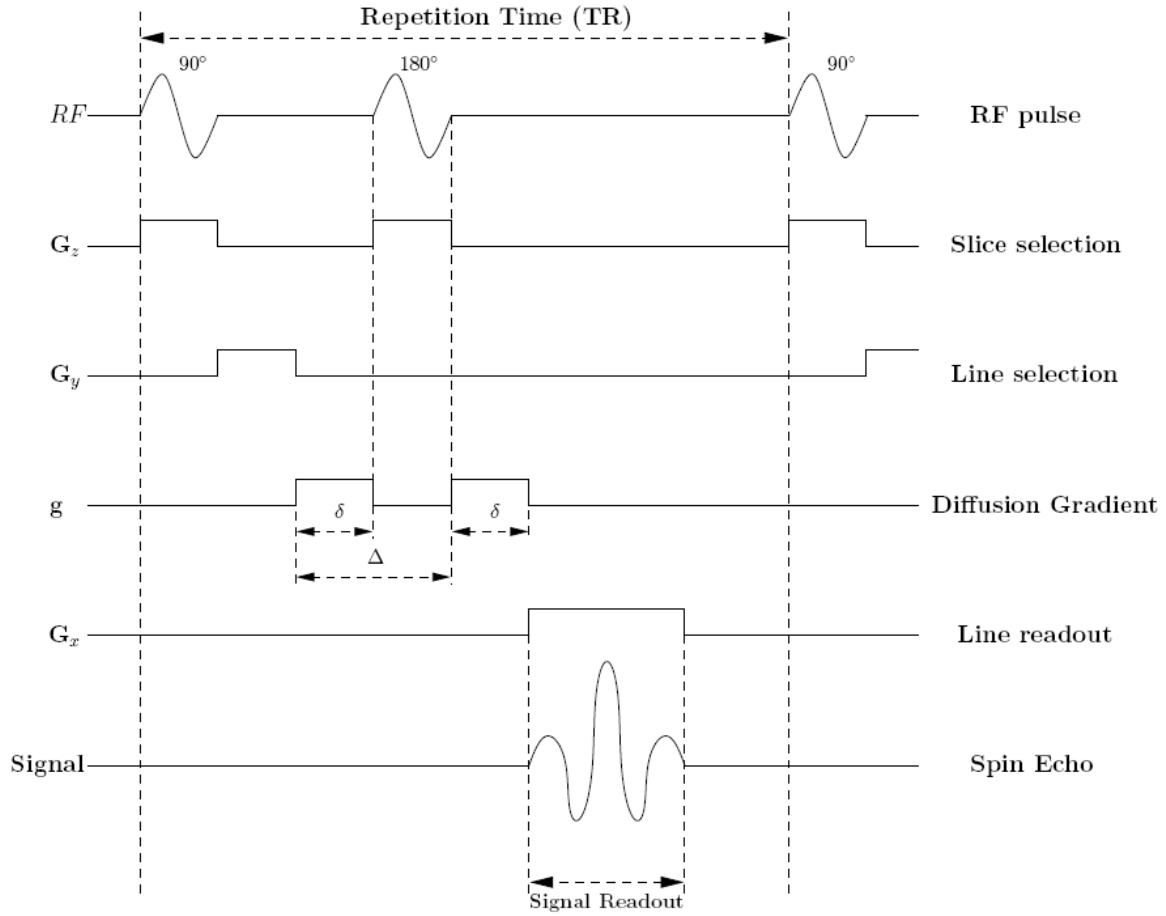


FIG. 2 – Séquence d'acquisition de Stejskal-Tanner (extrait de [6]).

2.3 Signal et bruit

Comme introduit plus haut, le signal mesuré est un courant complexe, créé par induction électromagnétique dans une bobine.

2.3.1 Caractérisation du bruit en IRM classique

Sijbers décrit dans sa thèse [5] que, étant données les sources du bruit en IRM (pertes resistives dans la chaîne d'acquisition et dans l'objet à imager, écrantage du champ radiofréquence par le corps humain...), il était justifié de considérer le bruit du signal complexe mesuré comme étant gaussien additif, décorrellé et de moyenne nulle. La transformée de Fourrier étant une opération linéaire et orthogonale, le signal complexe reconstruit est encore corrompu par un

bruit gaussien. Bien que la partie réelle et la partie imaginaire du signal comportent toute l'information d'intérêt, la représentation du signal par son amplitude, la phase étant très sensible aux irrégularités du champ. L'amplitude M est obtenue à partir des composantes réelle et imaginaire (R et I) par $M = \sqrt{R^2 + I^2}$, qui est une transformation non linéaire. Le signal M n'est plus gaussien, mais suit une distribution de Rice (*cf.* figure 3) :

$$p(M|A,\sigma) = \frac{M}{\sigma^2} \exp\left(-\frac{A^2 + M^2}{2\sigma^2}\right) I_0\left(\frac{AM}{\sigma^2}\right)$$

qui pour un faible *SNR* approche une distribution de Rayleigh, et pour un *SNR* élevé peut être approximée par une gaussienne. Sijbers *et al.* proposent une estimation au maximum de vraissemblance des paramètres d'une variable aléatoire Ricienne dans [7].

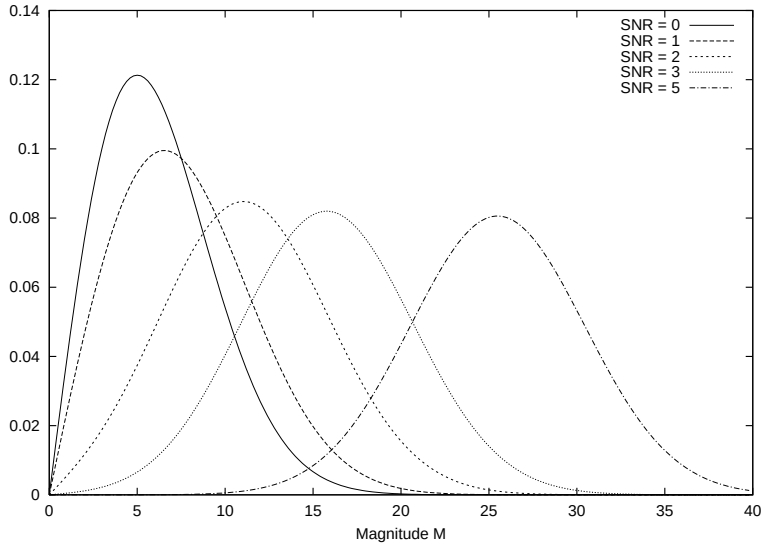


FIG. 3 – Distribution Ricienne pour différentes valeurs du *SNR*.

2.3.2 Application à l'IRM de diffusion

Plus encore qu'en IRM structurelle, l'IRM de diffusion exploite quantitativement les données acquises pour calculer et comparer l'anisotropie par exmple. De plus, les temps d'acquisitions sont limités par la pratique ainsi le *SNR* est faible et comme évoqué plus haut, la densité de probabilité du signal est significativement différente d'une gaussienne. Cette caractérisation du bruit est primordiale pour éviter d'introduire un biais lors de l'estimation des paramètres de diffusion. Dans ce sens, Basu *et al.* [8] étudient l'effet du bruit Ricien sur le tenseur de diffusion estimé. Il a tendance à biaiser positivement l'anisotropie fractionnelle et négativement la diffusivité moyenne du tenseur estimé. Par ailleurs ils montrent que le biais peut dépendre de l'orientation principale du tenseur par rapport aux gradients de diffusion.

Cette relation entre la direction des gradients de diffusion et la qualité de l'estimation sous-entend que le choix des directions et de l'amplitude des gradients de diffusion n'est pas sans conséquence sur la précision de la reconstruction. Deux études menées indépendemment par Papadakis *et al.* [9] et par Hasan *et al.* [10] montrent que la variance du tenseur reconstruit est directement liée aux choix des gradients de diffusion, en particulier pour les séquences à faible résolution angulaire (6 directions). Ils exhibent une séquence de six gradients (union de deux

FIG. 4 – Échantillonage de la sphère pour $N = 21$ et $N = 46$ directions.

bases othogonales) communément utilisée, mais cependant sous-optimale. Pour une description par un modèle d'ordre supérieur, plus de mesures sont nécessaires ; le choix de la séquence de gradients se fait par échantillonage approximativement uniforme de la sphère, la méthode de la tessellation permet d'obtenir des échantillonnages à $N = 5n^2 + 1, n = 1 \ldots \infty$ directions (voir figure 4).

Une étude menée par Basser et Pajevic [11] montre que la loi de distribution du tenseur estimé est normale, pourvu qu'on utilise un grand nombre de mesures (loi des grands nombres) et que le SNR ne soit pas trop faible. De cette observation, ils dérivent une évaluation de la fiabilité des caractéristiques calculées à partir du tenseur (valeurs propres, anisotropie...).

En plus du bruit additif, Mangin *et al.* [12] posent le problème des distortions géométriques échoplanaires. Elles sont introduites par les gradients de pondération en diffusion ; ainsi en se basant sur un modèle physique de distortion échoplanaires, ils proposent une estimation des paramètres de déformation, puis un recalage sur l'image acquise sans gradient. L'importance de ce recalage entre les différentes acquisition d'une même coupe semble primordial pour que les voxels de différentes images pondérées en diffusion d'une même coupe correspondent. On peut cependant remarquer que cette contribution est peu citée dans la littérature.

Ainsi bien que la nature du signal mesuré dans les images pondérées en diffusion est la même que pour l'IRM classique, il semble nécessaire de s'intéresser au problème du bruit spécifiquement pour cette modalité d'IRM. Par ailleurs, comme en traitement d'images classique, le SNR des images reconstruites peut être significativement augmenté en incluant des contraites *a priori* sur la solution recherchée. Nous présentons en dernière partie un panorama des algorithmes de reconstruction et de régularisation.

3 Estimation et régularisation

Bien que le thème du débruitage et de la régularisation ait été largement étudié en traitement du signal et de l'image, l'imagerie de diffusion y apporte de nouvelles problématiques. En effet, les modèles introduits en première partie sont complexes, et l'utilisation de la redondance spatiale pour les régulariser n'est pas triviale.

3.1 Reconstruction d'un champ de tenseur régulier

Malgré sa découverte récente, la littérature sur le tenseur de diffusion est vaste et la recherche reste très active. La première solution, apportée par Basser *et al.* [1] était une résolution aux moindres carrés des équations de Stejskal-Tanner. Bien que cela permette de bénéficier de la redondance apportée par un grand nombre d'acuqisitions, cela ne tenait pas compte de la contrainte de positivité pour le tenseur.

Plus tard, Tschumperlé et Deriche [13] proposent une approche variationnelle à base d'EDP pour minimiser une fonction de coût, dans laquelle un terme de régularisation appartaît. Le terme de régularisation utilisé dans la fonction de coût est une fonction croissante ϕ de la norme du gradient du tenseur :

$$\min_{\mathbf{D}\in\mathcal{P}(3)} \int_\Omega \sum_{k=1}^{n} \psi\left(\left|\ln\left(\frac{S(0)}{S(k)}\right) - \mathbf{g_k}^T\mathbf{D}\mathbf{g_k}\right|\right) + \alpha\phi\left(\|\nabla\mathbf{D}\|\right)$$

Le schéma numérique implémenté est une descente de gradient contrainte dans l'espace $\mathcal{P}(3)$ des matrices 3×3 symmétriques définies positives, suivant les travaux de la même équipe ([14]. Ainsi la solution est garantie d'appartenir à $\mathcal{P}(3)$. De plus, la régularisation du champ de tenseurs se fait conjointement à sa reconstruction.

Basu *et al.* [8] adoptent également une approche à base d'EDP, mais le terme d'attache aux données est remplacé par un terme de vraissemblance, qui intégre le modèle de bruit Ricien. La minimisation de la fonction de coût n'est pas contrainte, ainsi la solution peut ne pas être définie positive partout. Cependant l'intégration du modèle de bruit ricien est une contribution originale en IRM de diffusion, et montre des résultats prometteurs.

Fletcher [15] et Fillard [16] ont parallélement travaillé sur l'utilisation de la géométrie riemannienne, dans le but de trouver une représentation des tenseurs plus adaptée que l'espace euclidien. L'espace des tenseurs est alors vu comme une variété riemannienne, munie d'une métrique, ce qui permet de définir la notion de géodésique entre deux tenseurs, et ainsi de généraliser les opérations classiques dans les espaces euclidiens à la variété des tenseurs. Lenglet propose dans sa thèse [6] de redéfinir le gradient spatial de tenseurs grâce à la notion de géodésique. Le gradient calculé par cette technique tient compte de la géométrie de l'espace des tenseurs ; de plus il est invariant par changement d'échelle. Il serait intéressant de le substituer au gradient euclidien dans le critère proposé par Tschumperlé ; étant associé à une métrique invariante par changement d'échelle, ce gradient constitue un bon critère pour un filtrage anisotropique.

Revenons maintenant aux modèles d'ordre supérieur pour décrire la diffusion.

3.2 Modèles d'ordre supérieur

Comme introduit dans la première partie, il est possible de décrire le coefficient apparent de diffusion par une série harmonique sphérique. Pour le calcul des coefficients, Descoteaux *et al.* [4] offrent une alternative à la projection de la fonction discrète des valeurs mesurées dans l'espace des fonctions harmoniques. Ils proposent une fonction de coût aux moindres carrés, à laquelle ils ajoutent un terme de régularisation supplémentaire :

$$E(f) = \int_\Omega (\triangle_b f)^2 \,\mathrm{d}\Omega$$

qui peut, au vu de la définition de la base des fonctions harmoniques, s'exprimer très simplement en fonction des coefficients (on rappelle que $\Delta_b F = -l(l+1)F$). Dans cette étude, afin de pouvoir comparer les perfomances à d'autres approches similaires, Descoteaux explicite la

transformation (linéaire) entre le modèle des fonctions harmoniques sphériques et le tenseur d'ordre supérieur. Ceci permet également d'utiliser les mesures d'anisotropie définies sur les tenseurs d'ordre supérieur pour le modèle des fonction harmoniques sphériques.

Parmi les formalismes de modélisation pour le *HARDI*, on a également cité l'imagerie **q**-ball, en vue de calculer de manière approchée l'*odf* en appliquant le théorème de Funk-Radon. Partant de la représentation en harmoniques sphériques du signal sur la **q**-ball, Descoteaux *et al.* [17] dérivent une élégante méthode pour le calcul de la transformée de Funk-Radon, pour aboutir à une solution analytique et de faible complexité algorithmique. Le même schéma de régularisation à base de l'opérateur de Laplace-Beltrami permet de stabiliser l'estimation de la direction d'une fibre, et d'améliorer la détection des croisements de fibres.

Conclusion

En connaissant de manière approfondie les causes physiques du bruit en IRM, et en développant des techniques de régularisation adaptées, il a été possible de significativement diminuer l'influence du bruit dans les images de diffusion. À l'instar des derniers travaux de recherches sur le sujet, il nous semble nécessaire d'avancer vers une meilleure compréhension des mécanismes de la formation des images de diffusion, et des séquences d'acquisition qui leur sont associées.

Références

[1] P J Basser, J Mattiello, and D LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophys J*, 66(1) :259–267, Jan 1994.

[2] C A Clark and D Le Bihan. Water diffusion compartmentation and anisotropy at high b values in the human brain. *Magn Reson Med*, 44(6) :852–859, Dec 2000. Comparative Study.

[3] D C Alexander, G J Barker, and S R Arridge. Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magn Reson Med*, 48(2) :331–340, Aug 2002.

[4] Maxime Descoteaux, Elaine Angelino, Shaun Fitzgibbons, and Rachid Deriche. Apparent diffusion coefficients from high angular resolution diffusion imaging : estimation and applications. *Magn Reson Med*, 56(2) :395–410, Aug 2006.

[5] Jan Sijbers. *Signal and noise estimation from magnetic resonance images.* PhD thesis, Universiteit Antwerpen, October 1998.

[6] Christophe Lenglet. *Geometric and Variational Methods for Diffusion Tensor MRI Processing.* PhD thesis, Université de Nice-Sophia Antipolis, December 2006.

[7] J Sijbers, A J den Dekker, P Scheunders, and D Van Dyck. Maximum-likelihood estimation of Rician distribution parameters. *IEEE Trans Med Imaging*, 17(3) :357–361, Jun 1998.

[8] Saurav Basu, P. Thomas Fletcher, and Ross T. Whitaker. Rician noise removal in diffusion tensor mri. In *MICCAI (1)*, pages 117–125, 2006.

[9] N G Papadakis, D Xing, C L Huang, L D Hall, and T A Carpenter. A comparative study of acquisition schemes for diffusion tensor imaging using MRI. *J Magn Reson*, 137(1) :67–82, Mar 1999. Comparative Study.

[10] K M Hasan, D L Parker, and A L Alexander. Comparison of gradient encoding schemes for diffusion-tensor MRI. *J Magn Reson Imaging*, 13(5) :769–780, May 2001. Comparative Study.

[11] Peter J. Basser and Sinisa Pajevic. Dealing with uncertainty in diffusion tensor MR data. *Israel Journal of Chemistry*, 43 :129–144, 2003.

[12] Jean-Francois Mangin, C. Poupon, C. A. Clark, D. Le Bihan, and Isabelle Bloch. Eddy-current distortion correction and robust tensor estimation for mr diffusion imaging. In *MICCAI '01 : Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 186–194, London, UK, 2001. Springer-Verlag.

[13] David Tschumperlé and Rachid Deriche. Variational frameworks for dt-mri estimation, regularization and visualization. In *ICCV '03 : Proceedings of the Ninth IEEE International Conference on Computer Vision*, page 116, Washington, DC, USA, 2003. IEEE Computer Society.

[14] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras. Regularizing flows for constrained matrix-valued images. *J. Math. Imaging Vis.*, 20(1-2) :147–162, 2004.

[15] P Thomas Fletcher, Conglin Lu, Stephen M Pizer, and Sarang Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans Med Imaging*, 23(8) :995–1005, Aug 2004. Comparative Study.

[16] Pierre Fillard, Vincent Arsigny, Nicholas Ayache, and Xavier Pennec. A Riemannian framework for the processing of tensor-valued images. In Ole Fogh Olsen, Luc Florak, and Arjan Kuijper, editors, *Deep Structure, Singularities, and Computer Vision (DSSCV)*, number 3753 in LNCS, pages 112–123. Springer Verlag, June 2005.

[17] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. A fast and robust odf estimation algorithm in q-ball imaging. In *Biomedical Imaging : Macro to Nano, 2006. 3rd IEEE International Symposium on*, pages 81–84, 6-9 April 2006.