

Real-time Magnetic Resonance Q-Ball Imaging using Kalman Filtering with Laplace-Beltrami Regularization

Rachid Deriche^a and Jeff Calder^b

^aOdyssée Project Team, INRIA Sophia Antipolis - Méditerranée, France

^bDepartment of Mathematics and Statistics, Queen's University, Canada

ABSTRACT

Diffusion MRI has become an established research tool for the investigation of tissue structure and orientation from which has stemmed a number of variations, such as Diffusion Tensor Imaging (DTI) Diffusion Spectrum Imaging (DSI) and Q-Ball Imaging (QBI). The acquisition and analysis of such data is very challenging due to its complexity. Recently, an exciting new Kalman filtering framework has been proposed for DTI and QBI reconstructions in real time during the repetition time (TR) of the acquisition sequence. In this article, we first revisit and thoroughly analyze this approach and show it is actually *sub-optimal* and *not* recursively minimizing the intended criterion due to the Laplace-Beltrami regularization term. Then, we propose a new approach that implements the QBI reconstruction algorithm in real-time using a fast and robust Laplace-Beltrami regularization without sacrificing the optimality of the Kalman filter. We demonstrate that our method solves the correct minimization problem at each iteration and recursively provides the optimal QBI solution. We validate with real QBI data that our proposed real-time method is equivalent in terms of QBI estimation accuracy to the standard off-line processing techniques and outperforms the existing solution. This opens new and interesting opportunities for real-time feedback for clinicians during an acquisition and also for researchers investigating into optimal diffusion orientation sets and, real-time fiber tracking and connectivity mapping.

Keywords: High Angular Resolution Diffusion Imaging, Q-ball imaging, Orientation Distribution Function, Real-Time Diffusion MRI, Kalman filtering, Spherical Harmonics, Laplace-Beltrami Regularization

1. INTRODUCTION

Diffusion MRI (dMRI) is a recent Magnetic Resonance Imaging technique introduced in the middle of the 80's.¹⁻³ Since the first acquisitions of diffusion-weighted images (DWI) in vivo^{4,5} and the development of the rigorous formalism of the diffusion tensor (DT) model,⁶⁻⁸ dMRI has become an established research tool for the investigation of tissue structure and orientation and has opened up a landscape of extremely exciting discoveries for medicine and neuroscience. dMRI utilizes the measurement of Brownian motion of water molecules to gain information about tissue structure and orientation inside the brain and other organs. Using dMRI to infer the three dimensional diffusion probability displacement function (PDF) requires the acquisition of many diffusion images sensitized to different orientations in the sampling space. The number of diffusion weighted images (DWI) required depends on how the diffusion is modeled. The well known DT model assumes the PDF is Gaussian and requires at least 6 DWIs plus an additional unweighted image. However, the Gaussian assumption is an over-simplification of the diffusion of water molecules and thus has some limitations. While the Gaussian assumption is adequate for voxels in which there is only a single fiber orientation (or none), it breaks down for voxels in which there is more complicated internal structure. This is an important limitation, since resolution of DTI acquisition is between 1 mm³ and 27 mm³ while the physical diameter of fibers can be between 1 μ m and 30 μ m.^{9,10} Research groups currently agree that there is complex fiber architecture in most fiber regions of the brain.¹¹ In fact, it is currently thought that between one third to two thirds of imaging voxels in the human brain white matter contain multiple fiber bundle crossings.¹²

Therefore, it is of utmost importance to develop techniques that go beyond the limitations of diffusion tensor imaging (DTI). To do so, high angular resolution diffusion imaging (HARDI) has been proposed to measure diffusion images along several directions. Some HARDI reconstruction techniques are model dependent, some model-free, some have linear solutions whereas others require non-linear optimization schemes. For a thorough review of these methods the reader is

Email Addresses: Rachid.Deriche@sophia.inria.fr (Rachid Deriche), Jeff.W.Calder@gmail.com (Jeff Calder)

referred elsewhere.¹³ In this paper, we will be using a recent HARDI technique known as Q-Ball Imaging (QBI).¹⁴ Originally proposed by Tuch,¹⁴ QBI aims to reconstruct the angular portion of the diffusion displacement probability density function (PDF) of water molecules, also called the diffusion orientation distribution function (ODF). This is a spherical function that is particularly useful in tractography since its maxima are aligned with the underlying fiber directions at every voxel. QBI and the diffusion ODF play a central role in this work focused on the development of a real-time regularized ODF solution that outperforms the state-of-the-art ODF estimation. However, note that the regularized Kalman filtering framework could also be adapted easily to the problem of fiber orientation density (FOD)^{15,16} reconstruction.

Generally, HARDI acquisitions require many more diffusion weighted measurements than traditional DTI acquisitions, but they can resolve some of the fiber crossing problems. This comes at the price of a longer acquisition time, which can be problematic for clinical studies involving children and people inflicted with certain diseases. Excessive motion of the patient during the acquisition process can force an acquisition to be aborted or make the diffusion weighted images useless. Thus, one would like to make only as many acquisitions as is necessary. According to the literature, this number is likely to be somewhere between 50 and 200 diffusion weighted measurements but this is still an open question.

In recent literature,^{17,18} an algorithm has been proposed for real-time estimation of the diffusion tensor and the orientation distribution function (ODF) using the Kalman filtering framework. We should note that the idea of using the Kalman filter for DTI reconstruction was suggested earlier,⁷ however, to our knowledge it had not been developed nor implemented until very recently.^{17,18} The DT model, without any positivity constraints, is linear and easily fits into the Kalman filtering framework. However, the recently proposed¹⁹ fast and robust analytical ODF reconstruction algorithms include a regularization term to deal with poor signal to noise (SNR) ratios. This adds a term to the minimization problem which does not immediately fit into the standard Kalman filtering framework. In an attempt to include this regularization term in the Kalman filter, the existing algorithm^{17,18} is forced to make some adjustments to the reconstruction model. These adjustments make the Kalman filtering algorithm *sub-optimal* in terms of the Laplace-Beltrami criterion with substantially large errors occurring at the beginning of the acquisition sequence. This has significant implications for the intended applications of the real-time dMRI processing. As we would like to stop the acquisition as soon as the estimation has converged, a good estimation of the ODFs is highly desirable at the beginning of the acquisition and thus, the development of an optimal and incremental solution is important.

We propose a Kalman filtering solution that will correctly incorporate the Laplace-Beltrami regularization term¹⁹ into the filter's parameters without changing the ODF reconstruction model. The basic idea is to go back to the derivation of the Kalman filtering equations and include this regularization term. The surprising result is that *only* the initial covariance matrix needs to be modified to correctly implement the regularization term. We will show that our proposed algorithm yields *optimal* ODF estimations at each iteration (ie: continuously from the beginning to the end of the acquisition) and hence clearly provides an important added value over the elegant and original approach.^{17,18}

As background for the reader, we describe the Kalman filter in section 2.2. In section 3, we first thoroughly analyze the existing real-time algorithm^{17,18} and then describe our proposed recursive solution and prove that it is optimal with respect to the Laplace Beltrami regularization criterion developed in recent literature.¹⁹ In section 4 we present various experimental results verifying our claims before concluding the paper.

2. BACKGROUND

Before presenting the body of this paper, we will provide some background on Kalman filtering and ODF reconstruction techniques as well as clarify some notation.

2.1 Linear Observation Model

Throughout this paper we will use the following notation for the linear observation model.

$$y_i = \mathbf{C}_i \mathbf{x} + \epsilon_i \quad (1)$$

where y_i is a scalar observation at the discrete time step i , $\mathbf{x} \in \mathbb{R}^n$ is the state vector we wish to estimate, \mathbf{C}_i is the $1 \times n$ observation matrix and ϵ_i is a Gaussian random variable with zero mean and variance R_i . We will use the notation $E[\cdot]$ for the the expected value of a random variable or vector. Furthermore, we will assume that $E[\epsilon_i \epsilon_j] = 0$ for $i \neq j$. We will

frequently combine all the measurements up to time k into a linear system in order to compute the least squares estimation of the state \mathbf{x} . In this case, we will use the notation

$$\mathbf{y}_k = \mathbf{B}_k \mathbf{x} + \eta_k \quad (2)$$

where $\mathbf{y}_k = [y_1 \ y_2 \ \dots \ y_k]^T$ is a vector of observations, $\eta_k = [\epsilon_1 \ \epsilon_2 \ \dots \ \epsilon_k]^T$ is the vector of Gaussian random variables, and $\mathbf{B}_k = [\mathbf{C}_1^T \ \mathbf{C}_2^T \ \dots \ \mathbf{C}_k^T]^T$ is the $k \times n$ matrix containing all the observation matrices up to time k . Here, we use the T operator to denote matrix or vector transposition. Let us denote the covariance matrix of η_k by \mathbf{W}_k . Since the noise sequence is zero mean, this is written as $\mathbf{W}_k = E[\eta_k \eta_k^T]$. Since we are assuming that the components of the Gaussian noise vector are uncorrelated, \mathbf{W}_k is diagonal with diagonal elements given by $\text{Var}(\epsilon_j) = R_j$ for $j = 1 \dots k$.

We will denote by $\hat{\mathbf{x}}_k$ the minimum variance estimation of the state \mathbf{x} given all observations from time $i = 1 \dots k$. Thus, we can express $\hat{\mathbf{x}}_k$ as

$$\hat{\mathbf{x}}_k = \arg \min_{\mathbf{x} \in \mathbb{R}^n} (\mathbf{y}_k - \mathbf{B}_k \mathbf{x})^T \mathbf{W}_k^{-1} (\mathbf{y}_k - \mathbf{B}_k \mathbf{x}). \quad (3)$$

The solution to equation (3) is easily obtained by differentiating with respect to \mathbf{x} . We obtain

$$\hat{\mathbf{x}}_k = (\mathbf{B}_k^T \mathbf{W}_k^{-1} \mathbf{B}_k)^{-1} \mathbf{B}_k^T \mathbf{W}_k^{-1} \mathbf{y}_k \quad (4)$$

2.2 Kalman Filtering

The Kalman Filter is a linear optimal recursive estimator.²⁰ It was originally designed for linear dynamic systems but has since been successfully used for non-linear systems and problems involving state constraints. There are many excellent textbook references on Kalman Filtering.²¹ We recall the observation model equation (1)

$$y_i = \mathbf{C}_i \mathbf{x} + \epsilon_i.$$

As we record each new measurement, we gain an incremental amount of information about the state \mathbf{x} and by combining all the recorded measurements into a linear system, we can compute the minimum variance estimation after k observations $\hat{\mathbf{x}}_k$ using equation (4).

$$\hat{\mathbf{x}}_k = (\mathbf{B}_k^T \mathbf{W}_k^{-1} \mathbf{B}_k)^{-1} \mathbf{B}_k^T \mathbf{W}_k^{-1} \mathbf{y}_k$$

However, as we only gain an incremental amount of information about the state between observations y_{k-1} and y_k , it is reasonable to expect that the corresponding change in the state estimate between $\hat{\mathbf{x}}_{k-1}$ and $\hat{\mathbf{x}}_k$ is also incremental. The Kalman Filter allows us to quantify this incremental estimation correction and express $\hat{\mathbf{x}}_k$ in terms of $\hat{\mathbf{x}}_{k-1}$ and the new observation. The Kalman filtering equations are given below.

$$\begin{cases} \hat{\mathbf{x}}_0 = E[\mathbf{x}_0] \\ \mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T] \\ \mathbf{G}_k = \mathbf{P}_{k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k-1} \mathbf{C}_k^T + R_k)^{-1} \\ \mathbf{P}_k = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k-1} \\ \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k-1} + \mathbf{G}_k (y_k - \mathbf{C}_k \hat{\mathbf{x}}_{k-1}) \end{cases} \quad (5)$$

$\mathbf{z}_k = y_k - \mathbf{C}_k \hat{\mathbf{x}}_{k-1}$ is generally called the innovations sequence and \mathbf{G}_k is called the Kalman gain matrix. It can be shown that

$$\mathbf{P}_k = E[(\mathbf{x} - \hat{\mathbf{x}}_k)(\mathbf{x} - \hat{\mathbf{x}}_k)^T] \quad (6)$$

so \mathbf{P}_k is the covariance matrix of the estimation error at time k .

2.3 Q-Ball Linear Model

QBI has the advantage over diffusion tensor imaging of being model independent. In his seminal paper,¹⁴ Tuch showed that it was possible to reconstruct a smoothed version of the diffusion orientation distribution function (ODF), $\psi(\mathbf{g}) = \int_0^\infty P(\mathbf{g}r) dr$, directly from a single shell HARDI acquisition. The method involves viewing the HARDI acquisition signals as samples of a function on the sphere and taking the Funk-Radon transform (FRT) to arrive at the diffusion ODF. Reliable computation of the FRT has been an active research area. We will use the fast and robust analytical method with Laplace-Beltrami regularization¹⁹ as it performs well at low SNR and is well-suited for real-time applications.

This method relies on the decomposition of the spherical HARDI signal onto a spherical harmonic (SH) basis. Let Y_ℓ^m denote the SH of order ℓ and degree m ($m = -\ell, \dots, \ell$) in the standard basis and $Y_j(j(\ell, m) = (\ell^2 + \ell + 2)/2 + m)$ be the SH in the modified real and symmetric basis.^{19,22} We will use the notation $\tilde{\mathbf{x}} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n]^T$ for coefficients of the truncated spherical harmonic representation of the HARDI signal and $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ for the spherical harmonic coefficients of the ODF. The HARDI signal can now be expressed with the linear observation model

$$S(\theta_i, \phi_i) = \sum_{j=1}^n \tilde{x}_j Y_j(\theta_i, \phi_i) = \tilde{\mathbf{C}}_i \tilde{\mathbf{x}} \quad (7)$$

where $n = (L + 1)(L + 2)/2$ is the number of terms in the modified spherical harmonic basis of order L , $S(\theta_i, \phi_i)$ is the measured HARDI signal in the polar direction (θ_i, ϕ_i) and $\tilde{\mathbf{C}}_i = [Y_1(\theta_i, \phi_i) \ Y_2(\theta_i, \phi_i) \ \dots \ Y_n(\theta_i, \phi_i)]$. Since we have a linear observation model, given any number of measurements, we can decompose the HARDI signal onto the spherical harmonic basis by computing the least squares solution from equation (4). However, as diffusion weighted images are typically very noisy, Descoteaux et al²² have proposed a robust regularized estimation algorithm using the Laplace-Beltrami operator, which leads to the following regularized minimization equation,

$$M(\tilde{\mathbf{x}}) = \left(\tilde{\mathbf{y}}_k - \tilde{\mathbf{B}}_k \tilde{\mathbf{x}} \right)^T \left(\tilde{\mathbf{y}}_k - \tilde{\mathbf{B}}_k \tilde{\mathbf{x}} \right) + \lambda \tilde{\mathbf{x}}^T \tilde{\mathbf{L}} \tilde{\mathbf{x}}, \quad (8)$$

where $\tilde{\mathbf{B}}_k = [\tilde{\mathbf{C}}_1^T \ \tilde{\mathbf{C}}_2^T \ \dots \ \tilde{\mathbf{C}}_k^T]^T$ and $\tilde{\mathbf{y}}_k = [S(\theta_1, \phi_1) \ S(\theta_2, \phi_2) \ \dots \ S(\theta_k, \phi_k)]^T$ contains the HARDI measurements.

Now we are really interested in relating the ODF to the HARDI signal and as the spherical harmonic functions are eigenfunctions of the FRT, there is a simple relationship between the HARDI signal and the ODF in the spherical harmonic basis. The spherical harmonic coefficients of the ODF, \mathbf{x} , are given by $\mathbf{x} = \frac{1}{S_0} \mathbf{P} \tilde{\mathbf{x}}$ where \mathbf{P} is the $n \times n$ FRT matrix with diagonal elements p_{kk} given by $p_{kk} = 2\pi P_{\ell(j)}(0)$ and $P_{\ell(j)}(0)^*$ is the Legendre polynomial of degree ℓ evaluated at 0. Since we want to estimate the ODF directly from HARDI data, we can rewrite the minimization criterion (8) in terms of the spherical harmonic representation, \mathbf{x} , of the ODF. We have

$$M(\mathbf{x}) = (\mathbf{y}_k - \mathbf{B}_k \mathbf{x})^T (\mathbf{y}_k - \mathbf{B}_k \mathbf{x}) + \lambda \mathbf{x}^T \mathbf{L} \mathbf{x} \quad (9)$$

where $\mathbf{B}_k = \tilde{\mathbf{B}}_k \mathbf{P}^{-1}$ and $\mathbf{y}_k = \tilde{\mathbf{y}}_k / S_0$ and $\mathbf{L} = \mathbf{P}^{-1} \tilde{\mathbf{L}} \mathbf{P}^{-1}$. So by minimizing equation (9) we can robustly estimate the ODF directly from the HARDI signal. Now, although we have the same linear observation model as used in the Kalman filter, the minimization criterion is slightly different than equation (3) due to the additional regularization term. In the next section, we will address this problem in the Kalman filtering framework.

3. KALMAN FILTERING WITH REGULARIZATION

The Kalman filtering algorithm presented in section 2.2 recursively solves equation (3) at each time step k . However, we would like to consider the following minimization problem

$$\hat{\mathbf{x}}_k = \arg \min_{\mathbf{x} \in \mathbb{R}^n} (\mathbf{y}_k - \mathbf{B}_k \mathbf{x})^T \mathbf{W}_k^{-1} (\mathbf{y}_k - \mathbf{B}_k \mathbf{x}) + \lambda \mathbf{x}^T \mathbf{L} \mathbf{x} \quad (10)$$

which is slightly modified from equation (3) by the addition of $\lambda \mathbf{x}^T \mathbf{L} \mathbf{x}$ which is a quadratic regularization term on the state variable \mathbf{x} . This is the form of the minimization obtained using the robust Laplace Beltrami regularization described in the previous section. The closed form solution to equation (10) is easily obtained by setting the partial derivatives equal to zero. This yields

$$\hat{\mathbf{x}}_k = (\mathbf{B}_k^T \mathbf{W}_k^{-1} \mathbf{B}_k + \lambda \mathbf{L})^{-1} \mathbf{B}_k^T \mathbf{W}_k^{-1} \mathbf{y}_k \quad (11)$$

We would like to modify the Kalman filtering equations (5) so that we can recursively solve equation (11) with the addition of each new observation. Currently, we are only aware of one work^{17,18} addressing this problem in practice for DT and ODF estimation with the Kalman filtering framework, although the idea was mentioned earlier for DTI reconstruction.⁷ We show in the next sections that the existing method^{17,18} is actually *sub-optimal* for every time step except the last one. Then, in section 3.2, we present a new Kalman filtering method that deals with this regularization term with a very simple and elegant modification of the Kalman filtering equations from section 2.2.

* $\ell(j)$ is the order associated with the j^{th} element of the spherical harmonic basis, ie: for $j = 1, 2, 3, 4, 5, 6, 7, \dots$ $\ell(j) = 0, 2, 2, 2, 2, 2, 4, \dots$

3.1 Existing Method

The idea behind the recently proposed Kalman filtering solution^{17,18} is to start with the solution to equation (11) and “invert” it to obtain a linear system, whose least squares solution coincides with equation (11). Let N be the total number of measurements in the MR acquisition. The linear system obtained by “inverting” equation (11) is $\mathbf{y}_N = \mathbf{B}^+ \mathbf{x} + \eta_N$ where \mathbf{B}^+ is defined as

$$\mathbf{B}^+ = ((\mathbf{B}_N^T \mathbf{B}_N + \lambda \mathbf{L})^{-1} \mathbf{B}_N^T)^\dagger$$

The symbol $(\cdot)^\dagger$ is used to denote the Moore-Penrose pseudo-inverse operator. Now, denote the k^{th} row of \mathbf{B}^+ by \mathbf{C}_k^+ . By using the linear system $\mathbf{y}_N = \mathbf{B}^+ \mathbf{x} + \eta_N$, we would be using the following observation equation

$$y_k = \mathbf{C}_k^+ \mathbf{x} + \epsilon_k, \quad (12)$$

as input to the Kalman filter at each time step k . By construction, solving this system after all N measurements will yield the correct estimation, but there is no reason why this should be the correct incremental solver for any of the intermediate steps. In the next section, we analyze this approach and derive an explicit expression for the estimation at each iteration. In doing so, we show that this approach in fact *does not* minimize equation (10) at any iterations *except* for the last. Thus it is not the correct incremental solver and is *only* correct after the scan has been completed.

3.1.1 Analysis

In keeping with the proposed method,^{17,18} we assume the noise covariance matrix \mathbf{W}_k is the identity matrix for all time steps k . Before doing any detailed analysis, we can make some notes just from the definition of \mathbf{B}^+ and the observation equation

$$y_k = \mathbf{C}_k^+ \mathbf{x} + \epsilon_k. \quad (13)$$

As \mathbf{C}_k^+ is a row of \mathbf{B}^+ , it is clear from definition of \mathbf{B}^+ that every observation matrix \mathbf{C}_k^+ depends on \mathbf{B}_N and therefore on *all* N gradient acquisition directions. Hence, the acquisition of a given diffusion gradient orientation will be modeled differently depending on how many acquisitions are required in the acquisition. Furthermore, this means that when applying the Kalman filter to this model, every estimation depends on the entire diffusion orientation set being used. Thus, changing some diffusion gradients at the end of the acquisition could influence the estimations from the beginning. This is definitely undesirable behavior for an incremental optimal estimator. Based on experimental results,^{17,18} this method converges after all N acquisitions (as we mentioned before), but there are no results regarding how close the intermediate estimations are to the optimal intermediate estimations. This is the problem we will address next.

The best way to analyze this method is to derive an explicit expression for the Kalman filter output at each iteration under this method. First, we derive in other works²³ a closed form expression for \mathbf{B}^+ . The result is

$$\begin{aligned} \mathbf{B}^+ &= ((\mathbf{B}_N^T \mathbf{B}_N + \lambda \mathbf{L})^{-1} \mathbf{B}_N^T)^\dagger \\ &= \mathbf{B}_N \left(I_{n \times n} + \lambda (\mathbf{B}_N^T \mathbf{B}_N)^{-1} \mathbf{L} \right). \end{aligned} \quad (14)$$

Now let $\mathbf{D}_k = [\mathbf{C}_1^{+T} \ \mathbf{C}_2^{+T} \ \dots \ \mathbf{C}_k^{+T}]^T$ be the $k \times n$ matrix containing the first k observation matrices (or rows of \mathbf{B}^+). Using the result in equation (14), we can write \mathbf{D}_k as

$$\mathbf{D}_k = \mathbf{B}_k \left(I_{n \times n} + \lambda (\mathbf{B}_N^T \mathbf{B}_N)^{-1} \mathbf{L} \right) \quad (15)$$

This formula comes from the fact that taking the first k rows of the product of two matrices (ie: $\mathbf{D}_N = \mathbf{B}_{N \times n} \mathbf{A}_{n \times n}$) is equivalent to taking the first k rows of the first matrix, $\mathbf{B}_{k \times n}$ and multiplying them by the second matrix $\mathbf{A}_{n \times n}$ (ie: $\mathbf{D}_k = \mathbf{B}_{k \times n} \mathbf{A}_{n \times n}$). Now consider the linear system

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{x} + \eta_k. \quad (16)$$

This linear system relates all the observations up to time k to the state \mathbf{x} using the observation equation (12) from the proposed method. So, after k observations, the “optimal” state estimate will be given by the least squares solution to this linear system. That is, the output of the Kalman filter, $\hat{\mathbf{x}}_k$ at iteration k is given by

$$\hat{\mathbf{x}}_k = (\mathbf{D}_k^T \mathbf{D}_k)^{-1} \mathbf{D}_k^T \mathbf{y}_k$$

This can be simplified²³ to

$$\hat{\mathbf{x}}_k = \left(\mathbf{B}_k^T \mathbf{B}_k + \lambda \mathbf{B}_k^T \mathbf{B}_k (\mathbf{B}_N^T \mathbf{B}_N)^{-1} \mathbf{L} \right)^{-1} \mathbf{B}_k^T \mathbf{y}_k \quad (17)$$

A quick comparison with equation (11) shows that this is *not* the optimal estimate unless

$$\mathbf{B}_k^T \mathbf{B}_k = \mathbf{B}_N^T \mathbf{B}_N$$

or $\lambda = 0$. The former is only satisfied when $k = N$. Thus, we have shown that this method is sub-optimal at every estimation step except for the last. The implications of this sub-optimality should be made perfectly clear. If the expected number of acquisitions is set to $N = 200$ and the acquisition was stopped before completion (say at $k = 100$) the ODF estimates obtained from the Kalman filtering method recently proposed^{17,18} will *not* be minimizing the correct criterion from equation (10). We would in fact be minimizing

$$M(\mathbf{x}) = (\mathbf{y}_k - \mathbf{D}_k \mathbf{x})^T (\mathbf{y}_k - \mathbf{D}_k \mathbf{x}) \quad (18)$$

which yields a *different minimizing argument* (17) compared to the correct criterion (10) for $k < 200$.

3.1.2 Some Naive Approaches

There are some naive approaches one might take to correct this method. The first comes from the observation that this method could be "corrected" by replacing N by k in the definition of \mathbf{D}_k (15). Then the least squares solution to $\mathbf{y}_k = \mathbf{D}_k \mathbf{x}$ would be equal to the correct solution from equation (11). This in fact would be true, but it is impossible to apply Kalman filtering in this situation. Recall that the Kalman filter is derived from the linear observation model (1); it depends heavily on the fact that \mathbf{B}_{k-1} and \mathbf{B}_k differ by *only* the addition of one row \mathbf{C}_k . If we were to make this modification to \mathbf{D}_k , then \mathbf{D}_k would differ from \mathbf{D}_{k-1} by *much more* than just the addition of a single row. Therefore, there would be *no system of observation models* of the form $y_k = \mathbf{C}_k \mathbf{x} + \epsilon_k$ such that $\mathbf{D}_k = [\mathbf{C}_1^T \ \mathbf{C}_2^T \ \dots \ \mathbf{C}_k^T]^T$ for all k . This makes it impossible to derive a recursive relationship between $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{x}}_{k-1}$ and hence impossible to apply the Kalman filter.

The second approach is similar to the first; since we cannot apply the Kalman filter recursively with the modification described above, we could apply the Kalman filter to *all* of the data when we get a new observation. But by doing this, the solution is no longer obtained in a recursive way and real-time processing is impossible. For example, to compute the estimation at iteration k the Kalman filter would have to be iterated over all the previously recorded observations which is *k times more expensive* than the truly recursive Kalman filter and thus not suitable for real-time applications. In fact, there is little difference between this method and computing the offline estimation at each iteration.

The purpose of this discussion is to clarify that there are no simple modifications that could correct this method. In the next section, we will present a modification of the Kalman filter that allows the regularization term to be incorporated while still producing the optimal estimation at *every* time step.

3.2 Our Proposed Method

We propose a new Kalman filtering method that will incorporate the regularization term in the initial condition of \mathbf{P}_0 . Recall that we wish to solve the following equation (10)

$$\hat{\mathbf{x}}_k = \arg \min_{\mathbf{x} \in \mathbb{R}^n} (\mathbf{y}_k - \mathbf{B}_k \mathbf{x})^T \mathbf{W}_k^{-1} (\mathbf{y}_k - \mathbf{B}_k \mathbf{x}) + \lambda \mathbf{x}^T \mathbf{L} \mathbf{x}$$

in a recursive manner. The Kalman filtering equations for our proposed approach are

$$\left\{ \begin{array}{l} \hat{\mathbf{x}}_0 = E[\mathbf{x}_0] \\ \tilde{\mathbf{P}}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T (\mathbf{x}_0 - \hat{\mathbf{x}}_0)] \\ \mathbf{P}_0 = \left(\tilde{\mathbf{P}}_0^{-1} + \lambda \mathbf{L} \right)^{-1} \\ \mathbf{G}_k = \mathbf{P}_{k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k-1} \mathbf{C}_k^T + R_k)^{-1} \\ \mathbf{P}_k = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k-1} \\ \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k-1} + \mathbf{G}_k (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k-1}) \end{array} \right. \quad (19)$$

The state prediction correction and covariance recursive update equations are unchanged and the regularization term only appears in the initial condition. This is a significant result as it allows us to implement the regularization without any additional computational requirement compared to the standard Kalman filter. The derivation of these equations mimics the derivation of the standard Kalman filtering equations but is outside the scope of this paper. For more details on the derivation, please see other works by the authors.²³ With the correct Kalman filtering equations at hand we next present some experimental results that validate our proposed algorithm.

4. EXPERIMENTAL VALIDATION

In order to validate our method, we have tested it on real HARDI data and compared with the offline methods. Before we present the details of our results, we should clarify the exact algorithm that we have implemented as the Kalman filtering equations were derived in some generality.

4.1 Implementation Details

In any implementation of the Kalman filter, one must be careful in selecting proper initial conditions, most importantly so is the initial error covariance matrix \mathbf{P}_0 (written $\tilde{\mathbf{P}}_0$ in the previous section). One usually selects $\mathbf{P}_0 = \sigma^2 I_{n \times n}$, and chooses an appropriate value for σ . In the case of ODF estimation, before the first measurement, we have no *a priori* information about the diffusion properties at any voxel. To convey our ignorance of the system's state to the filter, we must set σ to a sufficiently large value so that the Kalman filter does not place any weight on the initial state \mathbf{x}_0 which we set to the zero vector. In our experiments we typically use $\sigma = 1000$.

The only other parameter to be set is the noise covariance matrix \mathbf{W}_k . To ensure that we are in fact minimizing equation (9), we use $\mathbf{W}_k = I_{k \times k}$ which means that $R_k = 1$ in all the Kalman filtering equations. The Kalman filtering equations used in our implementation are

$$\begin{cases} \hat{\mathbf{x}}_0 = 0 \\ \mathbf{P}_0 = (1/\sigma^2 I_{n \times n} + \lambda \mathbf{L})^{-1} \\ \mathbf{G}_k = \mathbf{P}_{k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k-1} \mathbf{C}_k^T + 1)^{-1} \\ \mathbf{P}_k = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k-1} \\ \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k-1} + \mathbf{G}_k (y_k - \mathbf{C}_k \hat{\mathbf{x}}_{k-1}) \end{cases}$$

where $\sigma = 1000$. We should note here that in the case where \mathbf{L} is invertible, we can take the limit as $\sigma \rightarrow \infty$ and initialize $\mathbf{P}_0 = (\lambda \mathbf{L})^{-1}$. However, for the case of ODF reconstruction, the Laplace-Beltrami operator has no dependence on the zeroth order spherical harmonic coefficient which makes \mathbf{L} singular.

4.2 Results and Discussion

Two sets of MRI data have been used in these experiments. In all experiments, a rank 4 spherical harmonic basis was used so there are 15 coefficients to estimate. The comparison metric used is the mean squared error (MSE) between the spherical harmonic coefficients of the ODFs over all relevant portions of the brain. The relevant regions of the brain were deduced from the T2 image. The first database is from the CEA Neurospin Lab in Paris, France.²⁴ The second is from the Max Planck Institute (MPI) in Leipzig, Germany.²⁵ We will refer to these by the acronyms CEA and MPI from now on. The CEA data was acquired on a 1.5T scanner with 200 encoding directions, $b = 3000 \text{ s/mm}^2$, 60 slices with 2mm thickness, 25 $b = 0 \text{ s/mm}^2$ images, 128x128 image matrix, TE=93.2 ms, TR = 19 s. The MPI data was acquired with a 3T scanner using 60 encoding directions and the average of three measurements per direction, seven $b = 0$ images, a b-value of 1000 s/mm^2 , 72 slices with 1.7 mm thickness, 128x128 image matrix, TE = 100 ms, TR = 12s.

The proposed algorithm has been tested and compared at each iteration with the optimal offline estimation. We have also implemented the only other existing method^{17,18} and tested it with the same datasets. Figure 1 shows the evolution of the MSE for the MPI and the CEA datasets for both our proposed regularized Kalman filtering algorithm and the existing algorithm proposed by Poupon et al. We can see that at the beginning of the acquisition, the MSE associated with the existing method is an order of magnitude higher than our proposed method. Our method yields a lower MSE for *every* iteration in the acquisition until both methods converge to the optimal offline estimation at the last iteration.

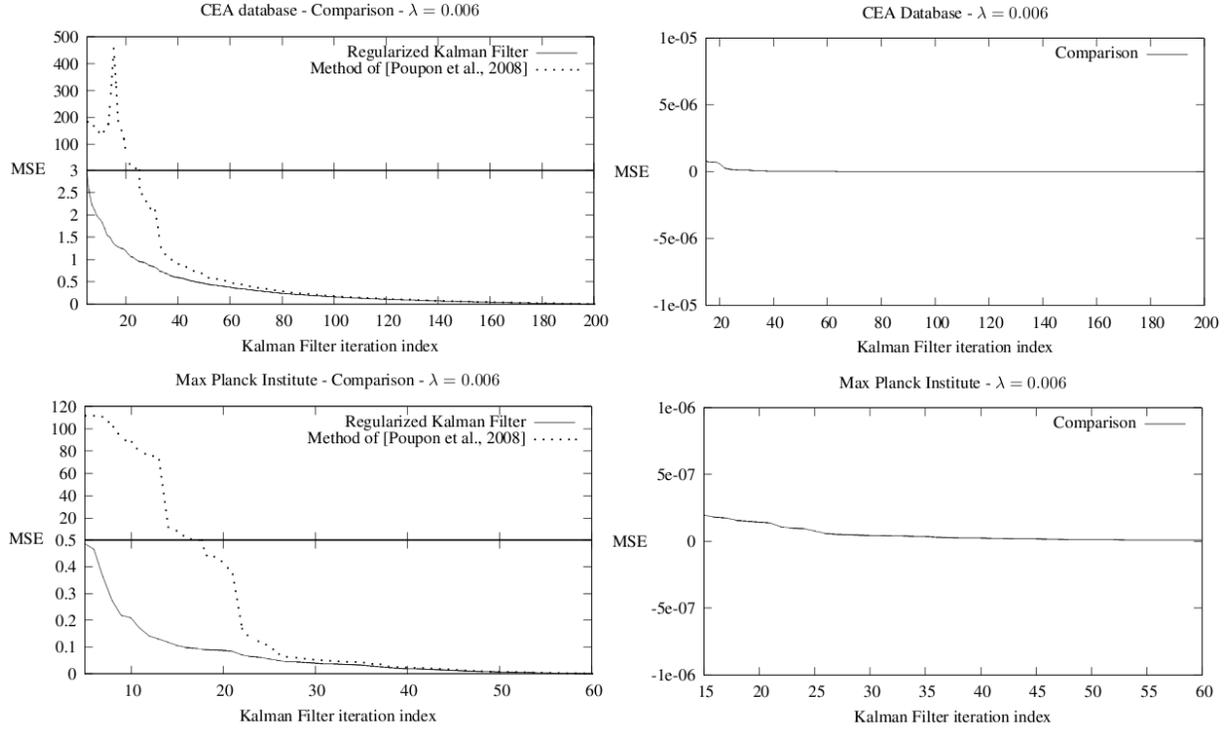


Figure 1. On the left, mean squared error between the SH coefficients of the final offline estimation and each iteration’s Kalman filter output using both our proposed Regularized Kalman filtering method and the existing method.¹⁸ This shows that our method is superior in terms of MSE at every iteration with the largest differences coming in the first half of the scan. Note that the top and bottom of the graphs on the left are on different scales to show all the important features. On the right, the MSE between each iteration’s optimal estimation and our Regularized Kalman filtering method. This shows that there is virtually no difference in terms of estimation accuracy between our method and the optimal off-line methods. We remind the reader that our algorithm has the advantage of recursively producing the results in real-time.

From these graphs we can conclude approximately how many measurements are actually necessary when using our proposed method. For the MPI dataset, there is little change between iteration 20 and 60, thus 20 measurements might be appropriate here. For the CEA dataset, the equivalent number is probably somewhere between 80 and 120. Keep in mind however, that these graphs are highly dependent on the fact that we have used a rank 4 spherical harmonic basis. Using a higher rank would result in more coefficients to be estimated and thus require more acquisitions. Furthermore, we cannot conclude with certainty how many measurements are required without investigating the effects on fiber tracking, fiber clustering and any other algorithms that depend on the ODF estimation.

The MSE comparisons on the left of figure 1 are very convincing, but they do not actually prove that our method yields the smallest MSE at each iteration. We do, however, have a closed form for the optimal estimation after each acquisition in equation (11), so this is what we should be comparing our solution to. We have computed offline the optimal estimation after each diffusion gradient orientation as per equation (11) and compared our Kalman filtering method to this in terms of MSE at each iteration. Our algorithm differs from the optimal by a maximum MSE of 10^{-6} throughout the entire scan. This difference is purely numerical and depends on our choice of $\sigma = 1000$. These results are shown on the right of figure 1. This validates our method by proving that we are in fact achieving the optimal estimation at each iteration.

To make a visual comparison between our proposed algorithm and the existing algorithm described in this paper^{17,18} we have generated ODF visualizations for a small region within a single axial slice of each dataset. The locations of the regions are shown in figure 2. Axial slice number 36 was chosen in both datasets and the ROIs were chosen in fiber crossing regions from the genu of the corpus callosum with peripheral fibers from the lateral cortex. In figure 3 we show the evolution of the ODFs under both algorithms next to the optimal estimation at that iteration for the MPI and CEA datasets respectively. We should note that the MPI HARDI data has a very high SNR, estimated to be roughly 37 (since

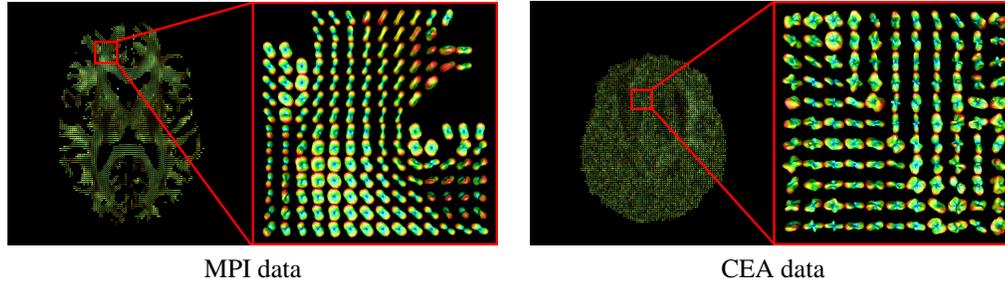


Figure 2. The Kalman filtering methods are visually compared in figure 3 on small regions of the MPI and CEA data respectively, illustrating fiber crossing regions from the genu of the corpus callosum with peripheral fibers from the lateral cortex.

3 averaging) in the white matter, so the visual difference between the optimal estimation using 15 versus 60 acquisitions is very small. Thus, this is an excellent example showing visually that the existing Kalman filtering method^{17,18} is sub-optimal. After 15 acquisitions, the optimal estimation is very good and close to the true ODFs, but the existing method is clearly quite far from optimal and in fact has very little in common visually with the true ODF field. After 20 acquisitions, the two methods are visually much more similar, but numerically still much different, and we see that by 60 acquisitions, they both converge to the optimal solution. For the CEA data, the ODFs become visually indistinguishable after about 60 to 80 acquisitions.

To validate that the algorithm is indeed real-time, we measured the execution time for the CEA dataset. The total processing time to update the Kalman filter across the entire volume (ie: all slices) after a new acquisition was approximately 6.23 seconds for both our algorithm and the existing method. As this is less than the repetition time of 12.5 seconds,^{17,18} both algorithms are truly real-time. The important difference between the algorithms is not the execution time, but that our proposed method solves the correct minimization problem at each iteration and hence provides a better estimation of the ODFs. For the test, both algorithms were implemented in C++ and executed on a 64-bit Linux machine with a dual-core 3.4GHz processor and 3.0 GBytes of RAM. However, the code was not written to take advantage of a multi-core processor and it was verified during the test that only one processor was being used. Theoretically, it would be possible to improve the performance by parallelizing the code, but this has not been investigated. We should note that we have defined real-time as the ability to produce the spherical harmonic coefficients of the ODF in real-time. In many applications, especially in a clinical setting, visualization is very important and one's definition of real-time depends on the applications at hand. Visualization is a very computationally intensive process and on a single workstation cannot be done in real-time. Further work could be aimed at implementing this algorithm on a cluster of workstations so visualization could also be performed in real-time.

Another topic for future work is the issue of the true Rician nature of the noise in the diffusion weighted images.²⁶⁻²⁸ In this work we have modeled the noise as Gaussian because it is required by the linear Kalman filter. At high SNR, the Gaussian distribution is a good approximation to Rician noise, but the approximation is poor at low SNR which is typical of diffusion MRI. We can see these differences in the quality of the estimations between the MPI and CEA datasets. The MPI data is acquired by averaging three measurements per orientation. Firstly, this lowers the noise variance, and secondly, from the central limit theorem in probability theory we know that the average of identically distributed random variables will tend to a Gaussian distribution. Thus, the noise in the MPI dataset is well modeled by a Gaussian distribution whereas the noise in the CEA data is not. The incorporation of the Rician noise model into a recursive filtering algorithm would likely yield better results for data with low SNR than the linear Kalman filter used in this work. However, this is a topic for future research.

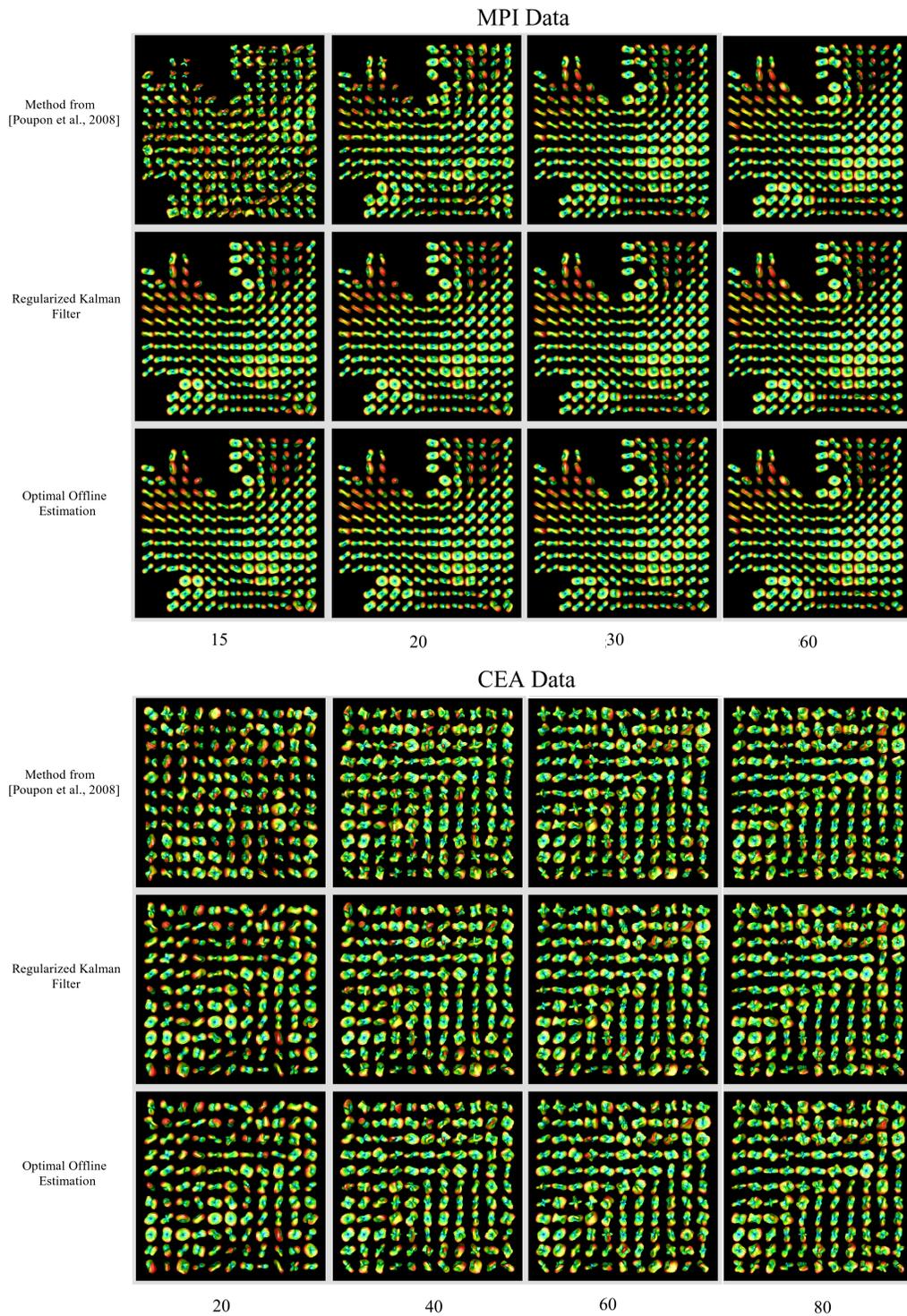


Figure 3. Visualization of the ODFs computed from the MPI (Top) and CEA (Bottom) data by (Top Row) Optimal least squares solution (Middle Row) our proposed Regularized Kalman filtering method, and (Bottom Row) the method proposed by^{17,18}. The ODFs are shown at intermediate diffusion acquisitions from left to right.

5. CONCLUSION

We have developed a truly incremental Regularized Kalman filtering algorithm for real-time processing of diffusion Magnetic Resonance Q-Ball Imaging. This method allows for real-time estimation of the orientation distribution function (ODF) during an ongoing MRI scan. We have validated that our proposed method is equivalent to the standard offline processing techniques and is therefore fit to be used as a real-time processing algorithm in MRI acquisitions. We have also analyzed the only other proposed method known to the authors^{17,18} and have shown that it is *sub-optimal* and not truly an incremental solution. Applications of real-time Kalman filtering for Q-ball imaging are numerous. It could be used to provide clinicians with a quality measure of the acquisition after each diffusion measurement and provide a stopping condition for the acquisition. This could shorten many HARDI acquisitions and reduce the probability of having to abort the acquisition due to motion of the patient. Aside from its obvious clinical applications, this real-time Kalman filtering framework will likely prove to be a useful tool for investigations into optimal diffusion gradient orientation sets, real-time fiber-tracking and connectivity mapping, real-time fiber orientation distribution function (FOD) estimation through spherical deconvolution,^{16,29} and many other areas. We are currently working towards achieving these objectives.

Acknowledgments

This work was partly supported by the ARC Diffusion MRI Program, the INRIA Internship Program, and the EADS Foundation. We are grateful and would like to express our sincere thanks to Maxime Descoteaux, Cyril Poupon, Fabrice Poupon and colleagues for their help in providing us the CEA Dataset and to Alfred Anwander and his colleagues for their help in providing us the MPI HARDI data.

REFERENCES

- [1] LeBihan, D. and Breton, E., "Imagerie de diffusion in vivo par résonance magnétique nucléaire," *C. R. Acad. Sci. Paris* **301 Série II**, 1109–1112 (1985).
- [2] Merboldt, K., Hanicke, W., and Frahm, J., "Self-diffusion nmr imaging using stimulated echoes," *J. Magn. Reson.* **64**, 479–486 (1985).
- [3] Taylor, D. and Bushell, M., "The spatial mapping of translational diffusion coefficients by the nmr imaging technique," *Physics in Medicine and Biology* **30**(4), 345–349 (1985).
- [4] Moseley, M., Cohen, Y., Mintorovitch, J., Kucharczyk, J., Tsuruda, J., Weinstein, P., and Norman, D., "Evidence of anisotropic self-diffusion," *Radiology* **176**, 439–445 (1990).
- [5] Osment, P., Packer, K., Taylor, M., Attard, J. J., Carpenter, T. A., Hall, L. D., Doran, S. J., and Herrod, N. J., "Nmr imaging of fluids in porous solids," *Phil. Trans. Roy. Soc.* **333**, 441–452 (1990).
- [6] Basser, P., Mattiello, J., Turner, R., and Bihan, D. L., "Diffusion tensor echo-planar imaging of human brain," in [*Proceedings of the SMRM*], 584 (1993).
- [7] Basser, P., Mattiello, J., and LeBihan, D., "MR diffusion tensor spectroscopy and imaging," *Biophysical Journal* **66**(1), 259–267 (1994).
- [8] Basser, P., Mattiello, J., and LeBihan, D., "Estimation of the effective self-diffusion tensor from the NMR spin echo," *Journal of Magnetic Resonance* **B**(103), 247–254 (1994).
- [9] Poupon, C., *Détection des faisceaux de fibres de la substance blanche pour l'étude de la connectivité anatomique cérébrale.*, PhD thesis, Ecole Nationale Supérieure des Télécommunications (Dec. 1999).
- [10] Beaulieu, C., "The basis of anisotropic water diffusion in the nervous system - a technical review," *NMR in Biomedicine* **15**, 435–455 (2002).
- [11] Pierpaoli, C., Jezzard, P., Basser, P., Barnett, A., and Chiro, G. D., "Diffusion Tensor MR imaging of human brain," *Radiology* **201**, 637–648 (1996).
- [12] Behrens, T. E. J., Johansen-Berg, H., Jbabdi, S., Rushworth, M. F. S., and Woolrich, M. W., "Probabilistic diffusion tractography with multiple fibre orientations. what can we gain?," *NeuroImage* **34**(1), 144–155 (2007).
- [13] Descoteaux, M., *High Angular Resolution Diffusion MRI: From Local Estimation to Segmentation and Tractography*, PhD thesis, Université de Nice - Sophia Antipolis (2008).
- [14] Tuch, D., "Q-ball imaging," *Magnetic Resonance in Medicine* **52**(6), 1358–1372 (2004).
- [15] Tournier, J.-D., Calamante, F., Gadian, D., and Connelly, A., "Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution," *NeuroImage* **23**, 1176–1185 (2004).

- [16] Tournier, J.-D., Calamante, F., and Connelly, A., “Robust determination of the fibre orientation distribution in diffusion mri: Non-negativity constrained super-resolved spherical deconvolution,” *NeuroImage* **35**(4), 1459–1472 (2007).
- [17] Poupon, C., Poupon, F., Roche, A., Cointepas, Y., Dubois, J., and Mangin, J. F., “Real-time mr diffusion tensor and q-ball imaging using kalman filtering.,” *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* **10**(Pt 1), 27–35 (2007).
- [18] Poupon, C., Roche, A., Dubois, J., Mangin, J.-F., and Poupon, F., “Real-time mr diffusion tensor and q-ball imaging using kalman filtering.,” *Med Image Anal* **12**, 527–34 (Jun 2008).
- [19] Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R., “Regularized, fast and robust analytical q-ball imaging.,” *Magnetic Resonance in Medicine* **58**, 497–510 (2007).
- [20] Kalman, R. E., “A new approach to linear filtering and prediction problems,” *Transactions of the ASME—Journal of Basic Engineering* **82**(Series D), 35–45 (1960).
- [21] Chui, C. K. and Chen, G., [*Kalman filtering with real-time applications*], Springer-Verlag New York, Inc., New York, NY, USA (1987).
- [22] Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R., “Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications,” *Magnetic Resonance in Medicine* **56**, 395–410 (2006).
- [23] Deriche, R., Calder, J., and Descoteaux, M., “Optimal real-time q-ball imaging using regularized kalman filtering with incremental orientation sets,” *INRIA Research Report* (2009).
- [24] Poupon, C., Poupon, F., Allirol, L., and Mangin, J.-F., “A database dedicated to anatomo-functional study of human brain connectivity,” in [*Twelfth Annual Meeting of the Organization for Human Brain Mapping (HBM)*], (2006).
- [25] Anwender, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., and Knosche, T. R., “Connectivity-based parcellation of broca’s area,” *Cerebral Cortex* **17**(4), 816–825 (2007).
- [26] Sijbers, J., den Dekker, A. J., Audekerke, J. V., Verhoye, M., and Dyck, D. V., “Estimation of the noise in magnitude mr images,” *Magnetic Resonance Imaging* **16**(1), 87–90 (1998).
- [27] Basu, S., Fletcher, T., and Whitaker, R., “Rician noise removal in diffusion tensor mri.,” *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* **9**(Pt 1), 117–125 (2006).
- [28] Fillard, P., Arsigny, V., Pennec, X., Hayashi, K. M., Thompson, P. M., and Ayache, N., “Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines,” *NeuroImage* **34**, 639–650 (2007).
- [29] Jian, B. and Vemuri, B. C., “A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted mri,” *IEEE Transactions on Medical Imaging* **26**(11), 1464–1471 (2007).