A Constrained Variational Principle for Direct Estimation and Smoothing of the Diffusion Tensor Field from DWI ⋆

Z. Wang[1], B.C. Vemuri[1], Y. Chen[2] and T. Mareci[3]

[1] Department of Computer & Information Sciences & Engr.,
University of Florida, Gainesville, FL 32611
[2] Department of Mathematics
University of Florida, Gainesville, FL 32611
[3] Department of Biochemistry
University of Florida, Gainesville, Fl. 32610

Abstract. In this paper, we present a novel constrained variational principle for simultaneous smoothing and estimation of the diffusion tensor field from diffusion weighted imaging (DWI). The constrained variational principle involves the minimization of a regularization term in an L^p norm, subject to a nonlinear inequality constraint on the data. The data term we employ is the original Stejskal-Tanner equation instead of the linearized version usually employed in literature. The original nonlinear form leads to a more accurate (when compared to the linearized form) estimated tensor field. The inequality constraint requires that the nonlinear least squares data term be bounded from above by a possibly known tolerance factor. Finally, in order to accommodate the positive definite constraint on the diffusion tensor, it is expressed in terms of cholesky factors and estimated. variational principle is solved using the augmented Lagrangian technique in conjunction with the limited memory quasi-Newton method. Both synthetic and real data experiments are shown to depict the performance of the tensor field estimation algorithm. Fiber tracts in a rat brain are then mapped using a particle system based visualization technique.

Index Terms - Diffusion Tensor MRI, Stjeskal-Tanner equation, nonlinear programing, augmented Lagrangian method, limited memory quasi-Newton.

1 Introduction

Diffusion is a process of movement of molecules as a result of random thermal agitation and in our context, refers specifi cally to the random translational motion of water molecules in the part of the anatomy being imaged with MR. In three dimension, water diffusivity can be described by a 3×3 matrix $\mathbf{D}$ called diffusion tensor which is intimately related to the geometry and organization of the microscopic environment.

⋆ This research was supported in part by the grant: NIH RO1-NS42075. Manuscript submitted to IPMI'03, also a Tech. Report, CISE-TR-03-004

General principle is that water diffuses preferably along ordered tissues e.g., the brain white matter.

Diffusion tensor MRI is a relatively new MR image modality from which anisotropy of water diffusion can be inferred quantitatively [2], thus provides a method to study the tissue microstructure e.g., white matter connectivity in the brain in vivo. Diffusion weighted echo intensity image S_l and the diffusion tensor $\mathbf{D}$ are related through the Stejskal-Tanner equation [2] as given by:

$$S_l = S_0 e^{-\mathbf{b_l}:\mathbf{D}} = S_0 e^{-\sum_{i=1}^{3}\sum_{j=1}^{3} b_{l,ij} D_{ij}} \tag{1}$$

where $\mathbf{b_l}$ is the diffusion weighting of the l-th magnetic gradient, ":" denotes the generalized inner product for matrices.

Taking $\log$ on both sides of equation (1) yields the following transformed linear Stejskal-Tanner equation:

$$log(S_l) = log(S_0) - \mathbf{b_l} : \mathbf{D} = log(S_0) - \sum_{i=1}^{3}\sum_{j=1}^{3} b_{l,ij} D_{ij} \tag{2}$$

Given several (at least seven) non-collinear diffusion weighted intensity measurements, $\mathbf{D}$ can be estimated via multivariate regression models from either of the above two equations. Diffusion anisotropy can then be computed to show microstructural and physiological features of tissues [3]. Especially in highly organized nerve tissue, like white matter, diffusion tensor provides a complete characterization of the restricted motion of water through the tissue that can be used to infer fiber tracts. The development of diffusion tensor acquisition, processing, and analysis methods provides the framework for creating fiber tract maps based on this complete diffusion tensor analysis [8, 11].

For automatic fiber tract mapping, the diffusion tensor field must be smoothed without losing relevant features. Currently there are two popular approaches, one involves smoothing the raw data S_l while preserving relevant detail and then estimate diffusion tensor $\mathbf{D}$ from the smoothed raw data (Parker et.al., [16, 23]). The raw data in this context consists of several diffusion weighted images acquired for varying magnetic field strengths and directions. Note that at least seven values at each 3D grid point in the data domain are required to estimate the six unknowns in the symmetric 2-tensor $\mathbf{D}$ and one scale parameter S_0. The raw data smoothing or de-noising can be formulated using variational principles which in turn requires solution to PDEs or at times directly using PDEs which are not necessarily arrived at from variational principles (see [17, 1, 24, 20, 7] and others in [6]).

Another approach to restore the diffusion tensor field is to smooth the principal diffusion direction after the diffusion tensor has been estimated from the raw noisy measurements S_l. In Poupon et al. [19], an energy function based on a Markovian model was used to regularize the noisy dominant eigenvector field computed directly from the noisy estimates of $\mathbf{D}$ obtained from the measurements S_l using the linearized Stejskal-Tanner equation (2). Coulon et.al., [9] proposed an iterative restoration scheme for principal diffusion direction based on direction map restoration work reported in [21]. Other sophisticated vector field restoration methods [22, 12, 18] can potentially be applied to the problem of restoring the dominant eigen-vector fields computed from the noisy estimates of $\mathbf{D}$. Recently, Chefd'Hotel et.al., [5] presented an elegant geometric solution to

the problem of smoothing a noisy $\mathbf{D}$ that was computed from S_l using the log-linearized model (2) described above. They assume that the given (computed) tensor field $\mathbf{D}$ from S_l is positive definite and develop a clever approach based on differential geometry of manifolds to achieve constrained smoothing where the smoothed tensor field is constrained to be positive semi-definite. Interesting results of mapped fibers are shown for human brain MRI.

We propose a novel formulation of the diffusion tensor field estimation and smoothing as a constrained optimization problem. The specific approach we use is called the augmented Lagrangian technique which allows one to deal with inequality constraints. *The novelty of our formulation lies in the ability to directly, in a single step process, estimate a smooth* $\mathbf{D}$ *from the noisy measurements* S_l *with the preservation of the positiveness constraint on* $\mathbf{D}$. *The formulation does not require any adhoc methods of setting parameter values to achieve the solution. These are the key features distinguishing our solution method from methods reported in literature to date.*

In contrast to our solution (to be described subsequently in detail), most of the earlier approaches used a two step method involving, (i) computation of a $\mathbf{D}$ from S_l using a linear least-squares approach and then (ii) computing a smoothed $\mathbf{D}$ via either smoothing of the eigen-values and eigen-vectors of $\mathbf{D}$ or using the matrix flows approach in [5]. The problem with the two step approach to computing $\mathbf{D}$ is that the estimated $\mathbf{D}$ in the first step using the log-linearized model need not be positive definite or even semi-definite. Moreover, it is hard to trust the fidelity of the eigen values and vectors computed from such matrices even if they are to be smoothed subsequently prior to mapping out the nerve fiber tracts. Also, the noise model used in the log-linearized scheme is not consistent with the physics.

Briefly, our model seeks to minimize a cost function involving, the sum of an L^p norm based gradient of the diffusion tensor $\mathbf{D}$ – whose positiveness is ensured via a Cholesky factorization to LL^T – and an L^p norm based gradient of S_0, subject to a nonlinear data constraint based on the original (not linearized) Stejskal-Tanner equation (1). The model is posed as a constrained variational principle which can be minimized by either discretizing the variational principle itself or the associated Euler-Lagrange equation. We choose the former and use the augmented Lagrangian method together with the limited memory quasi-Newton method to achieve the solution.

Rest of the paper is organized as follows: in section 2, the detailed variational formulation is described along with the nonlinear data constraints, the positive definite constraint and the augmented Lagrangian solution. Section 3 contains the detailed description of the discretization as well as the algorithmic description of the augmented Lagrangian framework. In section 4, we present experiments on application of our model to synthetic as well as real data. Synthetic data experiments are conducted to present comparison of tensor field restoration results with a recently presented work of Coulon et.al., [9]. Moreover, results of comparison between the use of the linearized Stejskal-Tanner model and the nonlinear form of the same are presented as well.

2 Constrained Variational Principle Formulation

Our solution to the recovery of a piecewise smooth diffusion tensor field from the measurements S_l is posed as a constrained variational principle. We seek to minimize a measure of lack of smoothness in the diffusion tensor $\mathbf{D}$ being estimated using the L^p norm of its gradient. This measure is then constrained by a nonlinear data fidelity term related to the original Stejskal-Tanner equation (1). This nonlinear data term is constrained by an inequality which requires that it be bounded from above by a possibly known tolerance factor. The positiveness constraint on the diffusion tensor being estimated is achieved via the use of the Cholesky factorization theorem from computational linear algebra. The constrained variational principle is discretized and posed using the augmented Lagrangian technique [14]. The augmented Lagrangian is then solved using the limited memory quasi-Newton scheme. The novelty of our formulation lies in the unified framework for recovering and smoothing of the tensor field from the data S. In addition, to our knowledge, this is the first formulation which allows for simultaneous estimation and smoothing of $\mathbf{D}$ as well as one in which the regularization parameter is not set in an adhoc way. The approach presented here describes a principled way to determine the regularization parameter.

Let $S_0(\mathbf{X})$ be the response intensity when no diffusion-encoding gradient is present, $\mathbf{D}(\mathbf{X})$ the unknown symmetric positive definite tensor, $\mathbf{L}\mathbf{L}^{\mathbf{T}}(\mathbf{X})$ be the Cholesky factorization of the diffusion tensor with $\mathbf{L}$ being a lower triangular matrix, $S_l, l = 1, .., N$ is the response intensity image measured after application of a magnetic gradient of known strength and direction and N is the total number of intensity images each corresponding to a direction and strength of the applied magnetic gradient.

$$\begin{aligned} \min \mathcal{E}(S_0, \mathbf{L}) &= \int_\Omega (|\nabla S_0(\mathbf{X})|^p + |\nabla \mathbf{L}(\mathbf{X})|^p) d\mathbf{X} \\ subject\ to \quad & \mathcal{C}(S_0, \mathbf{L}) = \alpha\sigma^2 - \int_\Omega \sum_{l=1}^{N} (S_l - S_0 e^{-\mathbf{b_l}:\mathbf{L}\mathbf{L}^{\mathbf{T}}})^2 d\mathbf{X} \geq 0 \end{aligned} \tag{3}$$

where Ω is the image domain, $\mathbf{b_l}$ is the diffusion weighting of the l-th magnetic gradient, ":" is the generalized inner product of matrices. The first and the second terms in the variational principle are L^p smoothness constraint on S_0 and $\mathbf{L}$ respectively, where $p > 12/7$ for S_0 and $p \geq 1$ for $\mathbf{L}$. $|\nabla \mathbf{L}|^p = \sum_d |\nabla L_d|^p$, where $d \in \mathcal{D} = \{xx, yy, zz, xy, yz, xz\}$ are indices to the six nonzero components of $\mathbf{L}$. The lower bounds on the value of p are chosen so as to make the proof of existence of a subproblem solution for this minimization (see section 2.4) mathematically tractable. α is a constant scale factor and σ^2 is the noise variance in the measurements S_l.

2.1 The Nonlinear Data Constraint

The Stejskal-Tanner equation (1) shows the relation between diffusion weighted echo intensity image S_l and the diffusion tensor $\mathbf{D}$. However, multivariate linear regression based on equation (2) has been used to estimate the diffusion tensor $\mathbf{D}$ [2]. It was pointed out in [2] that these results agree with nonlinear regression based on the

original Stejskal-Tanner equation (1). However, if the signal to noise ratio (SNR) is low and the number of intensity images S_l is not very large (unlike in [2] where $N = 315$ or $N = 294$), the result from multivariate linear regression will differ from the non-linear regression signifi cantly. A robust estimator belonging to the M-estimator family was used by Poupon et.al., [19], however, its performance is not discussed in detail. In Westin et. al., [25]), an analytical solution is derived from equation (2) by using a dual tensor basis, however, it should be noted that this can only be used for computing the tensor $\mathbf{D}$ when there is no noise in the measurements S_l or the SNR is extremely high.

Our aim is to provide an accurate estimation of diffusion tensor $\mathbf{D}$ for practical clinical use, where the SNR may not be high and the total number of intensity images N is restricted to a moderate number. The nonlinear data fi delity term based on original Stejskal-Tanner equation (1) is fully justifi ed for use in such situations. This nonlinear data term is part of an inequality constraint that imposes an upper bound on the closeness of the measurements S_l to the mathematical model $S_0 e^{-\mathbf{b_l}:\mathbf{L}\mathbf{L}^\mathbf{T}}$. The bound $\alpha\sigma^2$ may be estimated automatically from the measurements using any variance estimation methods from literature [13].

2.2 The L^p smoothness constraint

In Blomgren et.al., [4], it was shown that L^p smoothness constraint doesn't admit discontinuous solutions as the TV-norm does when $p > 1$. However, when p is chosen close to 1, its behavior is close to the TV-norm for restoring edges. In our constrained model, we need $p > 12/7$ for regularizing S_0 and $p >= 1$ for $\mathbf{L}$ to ensure existence of the solution. Note that what is of importance here is the estimation the diffusion tensor $\mathbf{D}$ and therefore, the edge-preserving property in the estimation process is more relevant for the case of $\mathbf{D}$ than for S_0. Hence, we choose an appropriate p for S_0 and $\mathbf{D}$ permitted by the theorem below. In our experiment, we choose $p = 1.705$ for S_0 and $p = 1.00$ for $\mathbf{L}$.

2.3 The Positive Definite Constraint

In general, a matrix $\mathbf{A} \in \Re^{n\times n}$ is said to be positive defi nite if $\mathbf{x}^\mathbf{T}\mathbf{A}\mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$ in $\Re^n$. The diffusion tensor $\mathbf{D}$ happens to be a positive defi nite matrix but due to the noise in the data S_l, it is hard to recover a $\mathbf{D}$ that retains this property unless one includes it explicitly as a constraint. One way to impose this constraint is using the Cholesky factorization theorem, which states that: *If* $\mathbf{A}$ *is a symmetric positive definite matrix then, there exists a unique factorization* $\mathbf{A} = \mathbf{L}\mathbf{L}^\mathbf{T}$ *where,* $\mathbf{L}$*is a lower triangular matrix with positive diagonal elements.* After doing the Cholesky factorization, we have transfered the inequality constraint on the matrix $\mathbf{D}$ to an inequality constraint on the diagonal elements of L. This is however still hard to satisfy theoretically because, the set on which the minimization takes place is an open set. However, in practise, with fi nite precision arithmetic, testing for a positive defi niteness constraint is equivalent to testing for positive semi-defi niteness. To answer the question of positive semi-defi niteness, a stable method would yield a positive response even for nearby symmetric matrices. This is because, $\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{E}$ with $\| \mathbf{E} \| \leq \epsilon \| \mathbf{D} \|$, where ϵ is a small multiple of

the machine precision. Because, with an arbitrarily small perturbation, a semi-definite matrix can become definite, it follows that in finite precision arithmetic, testing for definiteness is equivalent to testing for semi-definiteness. Thus, we repose the positive definiteness constraint on the diffusion tensor matrix as, $\mathbf{x^T D x} \geq 0$ which is satisfied when $\mathbf{D} = \mathbf{L L^T}$.

2.4 Comments on Existence of the Solution

Consider the augmented Lagrangian formulation which serves as a subproblem of (3):

$$\min_{(S_0,\mathbf{L})\in\mathcal{A}} \mathcal{L}(S_0,\mathbf{L};\lambda,\mu) = \mathcal{E}(S_0,\mathbf{L}) - \lambda\mathcal{C}(S_0,\mathbf{L}) + \frac{1}{2\mu}\mathcal{C}^2(S_0,\mathbf{L}) \tag{4}$$

Where $\mathcal{A} = \{(S_0,\mathbf{L}) \mid \mathbf{L} \in BV(\Omega),\ L_d \in L^2(\Omega), d \in \mathcal{D}\ and\ S_0 \in W^{1,p}(\Omega), p > 12/7\}$. Here $\Omega \subset \Re^3$, $BV(\Omega)$ denotes the space of bounded variation functions on the domain Ω, $L^2(\Omega)$ is the space of square integrable functions on Ω and $W^{1,p}(\Omega)$ denotes the Sobolev space of order p on Ω [10].

Theorem 1 *Suppose $S_l \in L^4(\Omega)$, then the augmented Lagrangian formulation (4) has a solution $(S_0,\mathbf{L}) \in \mathcal{A}$.*

Proof Outline:
We can prove the following:

- Lower semi-continuity of the first term $\mathcal{E}(S_0,\mathbf{L})$ in (4).
- Continuity of the second term $\mathcal{C}(S_0,\mathbf{L})$ in (4) for $S_0 \in W^{1,p}(\Omega)$ when $p > 6/5$.
- Continuity of the third term $\mathcal{C}^2(S_0,\mathbf{L})$ in (4) for $S_0 \in W^{1,p}(\Omega)$ when $p > 12/7$.

Thus, if $(S_0^{(n)},\mathbf{L}^{(n)})$ is a minimizing sequence, then it has a subsequence $(S_0^{(nk)},\mathbf{L}^{(nk)})$ where $\mathbf{L}^{(nk)})$ converges weakly in $BV(\Omega)$ and $S_0^{(nk)}$ converges weakly in $W^{1,p}(\Omega)$ when $p > 12/7$. From the compact embedding theorem [10], $\mathbf{L}^{(nk)}$ converges strongly in L^2 a.e. (almost everywhere) on Ω. Similarly, $S_0^{(nk)}$ converges strongly in L^4 a.e. on Ω. Thus the minimizing sequence has a convergent subsequence, and the convergence is the solution of the minimization problem (4)([10]).

Finding a solution of the constrained variation principle (3) involves solving a sequence of (4) with fixed λ and μ at each stage. It is much more difficult than when dealing with the problems of recovering and smoothing separately. However, there are benefits of posing the problem in this constrained unified framework, namely, one does not accumulate the errors from a two stage process. Moreover, this framework incorporates the nonlinear data term which is more appropriate for low SNR values prevalent when b is high. Also, the noise model is correct for the nonlinear data model unlike the log-linearized case. Lastly, in the constrained formulation, it is now possible to pose mathematical questions of existence and uniqueness of the solution – which was not possible in earlier formulations reported in literature.

3 Numerical Methods

We discretize the constrained variational principle (3) and then transform it into a sequence of unconstrained problems by using the augmented Lagrangian method and then employ the limited quasi-Newton technique [14] to solve them. Let

$$\begin{aligned} R_{l,ijk} &= S_{l,ijk} - S_{0,ijk} e^{-\mathbf{b_l}:\mathbf{L_{ijk}}\mathbf{L_{ijk}^T}} \\ |\nabla S_0|_{ijk} &= \left[\sqrt{(\Delta_x^+ S_0)^2 + (\Delta_y^+ S_0)^2 + (\Delta_z^+ S_0)^2 + \epsilon}\right]_{ijk} \\ |\nabla L_d|_{ijk} &= \left[\sqrt{(\Delta_x^+ L_d)^2 + (\Delta_y^+ L_d)^2 + (\Delta_z^+ L_d)^2 + \epsilon}\right]_{ijk}, \quad d \in \mathcal{D} \\ |\nabla \mathbf{L}|_{ijk}^p &= \sum_{d \in \mathcal{D}} |\nabla L_d|_{ijk}^p \end{aligned} \tag{5}$$

Where Δ_x^+, Δ_y^+ and Δ_z^+ are forward difference operators, ϵ is a small positive number used to avoid singularities of the L^p norm when $p < 2$. Now the discretized constrained variational principle can be written as:

$$\begin{aligned} &\min_{S_0, \mathbf{L}} \mathcal{E}(S_0, \mathbf{L}) = \sum_{i,j,k} (|\nabla S_0|_{ijk}^p + |\nabla \mathbf{L}|_{ijk}^p) \\ &\quad subject\ to \quad \mathcal{C}(S_0, \mathbf{L}) = \alpha\sigma^2 - \sum_{i,j,k} \sum_{l=1}^{N} R_{l,ijk}^2 \geq 0 \end{aligned} \tag{6}$$

The above problem is now posed using the augmented Lagrangian method, where a sequence of related unconstrained subproblems are solved, and the limit of these solutions is the solution to (6). Following the description in [14], the k-th subproblem of (6) is given by:

$$\min \mathcal{L}(S_0, \mathbf{L}, s; \lambda_k, \mu_k) = \mathcal{E}(S_0, \mathbf{L}) - \lambda_k(\mathcal{C}(S_0, \mathbf{L}) - s) + \frac{1}{2\mu_k}(\mathcal{C}(S_0, \mathbf{L}) - s)^2 \tag{7}$$

where $s \geq 0$ is a slack variable, μ_k, λ_k are the barrier parameter and the Lagrange multiplier for the k-th subproblem respectively.

One can explicitly compute the slack variable s at the minimum from

$$s = max((\mathcal{C}(S_0, \mathbf{L}) - \mu_k \lambda_k, 0)$$

and substitute it in (7) to get an equivalent subproblem in $(S_0, \mathbf{L})$ given by:

$$\begin{aligned} &\min \mathcal{L}_{\mathcal{A}}(S_0, \mathbf{L}; \lambda_k, \mu_k) \\ &= \begin{cases} \mathcal{E}(S_0, \mathbf{L}) - \lambda_k \mathcal{C}(S_0, \mathbf{L}) + \frac{1}{2\mu_k}\mathcal{C}^2(S_0, \mathbf{L}) & if\ \mathcal{C}(S_0, \mathbf{L}) - \mu_k\lambda_k \leq 0 \\ \mathcal{E}(S_0, \mathbf{L}) - \frac{\mu_k}{2}\lambda_k^2 & otherwise \end{cases} \end{aligned} \tag{8}$$

The following algorithm summarizes the procedure to find the solution for (6):

Initialize $S_0(0)$, $\mathbf{L}(0)$ using the nonlinear regression, choose initial μ_0 and λ_0.
for k = 1, 2, ...
Find approximate minimizer $S_0(k)$, $\mathbf{L}(k)$ of $\mathcal{L}_A(\cdot,\cdot;\lambda_k,\mu_k)$ starting with $S_0(k-1)$, $\mathbf{L}(k-1)$;
If final convergence test is satisfied
STOP with approximate solution $S_0(k)$, $\mathbf{L}(k)$;
Update Lagrange multiplier using $\lambda_{k+1} = max(\lambda_k - \mathcal{C}(S_0, \mathbf{L})/\mu_k, 0)$;
Choose new penalty parameter $\mu_{k+1} = \mu_k/2$;
Set new starting point for the next iteration to $S_0(k)$, $\mathbf{L}(k)$;
endfor

Due to the large number of unknown variables in the minimization, we solve the subproblem using limited memory Quasi-Newton technique. Quasi-Newton like methods compute the approximate Hessian matrix at each iteration of the optimization by using only the first derivative information. In Limited-Memory Broyden-Fletcher-Goldfarb-Shano (BFGS), search direction is computed without storing the approximated Hessian matrix. Details can be found in Nocedal et.al., [14].

4 Experimental Results

In this section, we present two sets of experiments on the application of our smoothing tensor estimation model. One is on synthetic data sets and the other is on a real data set consisting of a DWI acquired from a normal rat brain.

We synthesized an anisotropic tensor field on a 3D lattice of size 32x32x8. The volume consists of two homogeneous regions with the following values for S_0 and $\mathbf{D}$:

$$Region1: S_0 = 10.00 \quad \mathbf{D} = 0.001 \times [0.9697 \;\; 1.7513 \;\; 0.8423 \;\; 0.0 \;\; 0.0 \;\; 0.0]$$
$$Region2: S_0 = 8.33 \quad \mathbf{D} = 0.001 \times [1.5559 \;\; 1.1651 \;\; 0.8423 \;\; 0.3384 \;\; 0.0 \;\; 0.0]$$

Where the tensor $\mathbf{D}$ is depicted as $[d_{xx}, d_{yy}, d_{zz}, d_{xy}, d_{xz}, d_{yz}]$, the dominant eigen vector of the first region is along the y axis, while that of the second region is in the xy plane and inclined at 60 degrees to the y axis.

The diffusion weighted images S_l are generated using the Stejskal-Tanner equation at each voxel $\mathbf{X}$ given by:

$$S_l(\mathbf{X}) = S_0(\mathbf{X}) e^{-\mathbf{b_l}:\mathbf{D}(\mathbf{X})} + n(\mathbf{X}), \quad n(\mathbf{X}) \sim N(0, \sigma_N) \tag{9}$$

where $N(0, \sigma_N)$ is a zero mean Gaussian noise with standard deviation σ_N. We choose the 7 commonly used gradient configurations ([25]) and 3 different field strengths in each direction for $\mathbf{b_l}$ values.

Figure 1 shows the results for a synthetic data set with $\sigma_N = 1.5$. We display the dominant eigen vectors computed from the original and the restored diffusion tensor field using the following methods: (i) **Linear** - linear regression from (2) as in [2], (ii) **Nonlinear** - nonlinear regression from (1), (iii) **Linear + EVS** (Eigen vector smoothing) - linear regression followed by the dominant eigen vector smoothing method described in Coulon et.al., [9], (iv) **Nonlinear + EVS** - nonlinear regression plus the

smoothing as in (iii), and (v) **Our method**. It is evident from this figure that our new model yields very good estimates of the dominant eigen vector field. The method in Coulon et.al., however, will not work well at voxels where the estimated dominant eigen vectors are almost orthogonal to the ones in their neighborhoods. In such cases, Coulon et.al.'s method will treat them as discontinuities and doesn't smooth them. Though it is possible to treat these locations as outliers in Coulon et.al.'s method, it is difficult to set a reasonable criteria. Further quantitative measures (described below) also shows the superiority of our model.

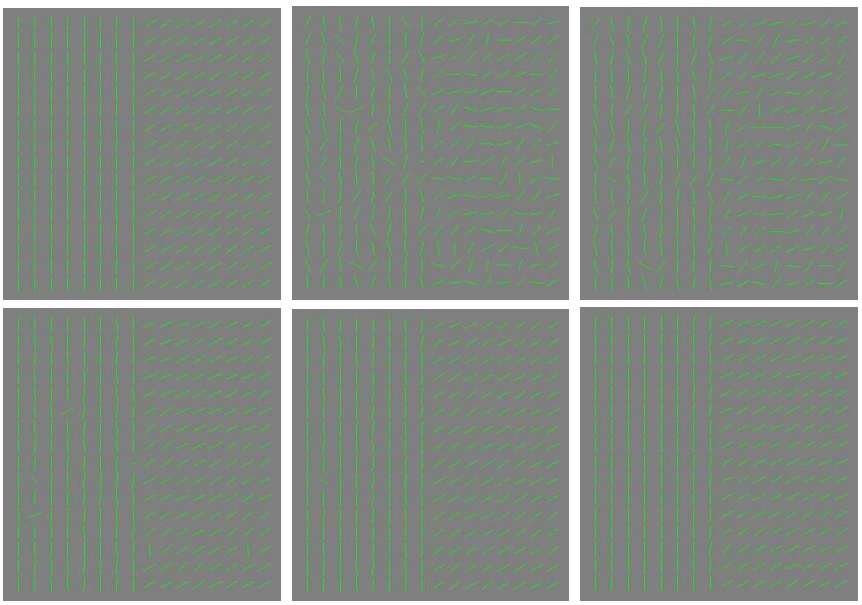


Fig. 1. A slice of the results for the synthetic data with $\sigma_N = 1.5$: Top left image is the dominant eigen vector computed from original tensor field, and the other images arranged from left to right, top to bottom are the dominant eigen vectors computed from estimated tensor field using the following methods: linear regression, nonlinear regression, linear + EVS, nonlinear+EVS and our model.

To quantitatively assess the proposed model, we compare the accuracy of the dominant eigen vector computed from previously mentioned methods. Let θ be the angle(in degrees) between the dominant eigen vector of the estimated diffusion tensor field and the dominant eigen vector of the original tensor field, table 1 shows the mean (μ_θ) and standard deviation(σ_θ) of θ using different methods for the synthetic data with different levels of additive Gaussian noises. A better method is one that yields smaller values. From this table, we can see the our model yields lower values than all other methods under various noise levels. It is also clear from this table that methods using the original nonlinear Stejskal-Tanner equation (1) are more accurate than those using the linearized

one (2). The advantage of our method and the nonlinear approaches are more apparent when the noise level is higher, which supports our discussion in section 2.1.

$\sigma_n = 0.5$

	Linear	Nonlinear	Linear+EVS	Nonlinear+EVS	Our method
μ_θ	10.00	8.22	1.76	1.46	0.76
σ_θ	7.29	5.90	2.38	1.44	1.17

$\sigma_n = 1.0$

	Linear	Nonlinear	Linear+EVS	Nonlinear+EVS	Our method
μ_θ	22.69	18.85	6.87	4.75	2.19
σ_θ	17.77	15.15	14.59	10.64	2.52

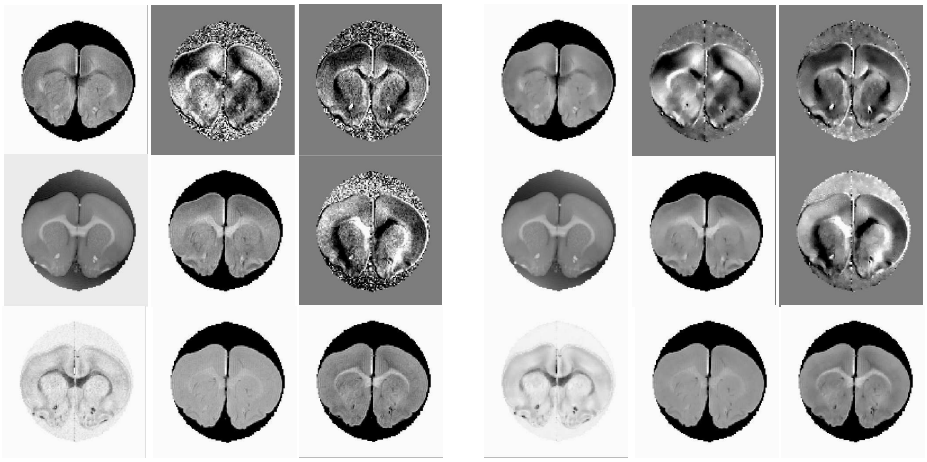
$\sigma_n = 0.5$

	Linear	Nonlinear	Linear+EVS	Nonlinear+EVS	Our method
μ_θ	34.10	30.26	16.29	12.50	6.47
σ_θ	23.11	22.22	24.37	22.19	9.58

Table 1. Comparison of the accuracy of the estimated dominant eigen vectors using different methods for different noise levels.

The normal rat brain data we used here has 21 diffusion weighted images measured using the same configuration of $\mathbf{b}_i$ as in the previous example, each image is a 128x128x78 volume data. We extract 10 slices in the region of interest, namely the corpus callosum, for our experiment. Figure 2(b) depicts images of the six independent components of the estimated diffusion tensor, the computed FA, the trace($\mathbf{D}$) and S_0 (echo intensity without applied gradient) obtained using our proposed model. As a comparison, figure 2(a) shows the same images computed using linear least square fitting based on the linearized Stejskal-Tanner equation from the raw data. For display purposes, we use the same brightness and contrast enhancement for displaying the corresponding images in the two figures. The effectiveness of edge preservation in our method is clearly evident in the off-diagonal components of $\mathbf{D}$. In addition, fiber tracts were estimated as integral curves of the dominant eigen vector field of the estimated $\mathbf{D}$ and is visualized using the particle systems technique (Pang et.al [15]). The mapped fiber tracts are found to follow the expected tracts quite well from a neuroanatomical perspective as shown in figure 3.

In the above presented results, what is to be noted is that we have demonstrated a proof of concept for the proposed simultaneous recovery and smoothing of $\mathbf{D}$ in the case of the synthetic data and the normal rat brain respectively. The quality of results obtained for the normal rat brain is reasonably satisfactory for visual inspection purposes, however intensive quantitative validation of the mapped fibers needs to be performed and will be the focus of our future efforts.



(a) (b)

Fig. 2. (a)Results of normal rat brain estimated using multivariate linear regression without smoothing. (b) Results of normal rat brain estimated using our proposed method. Both (a) and (b) are arranged as following: First row, left to right: D_{xx}, D_{xy} and D_{xz}. Second row, left to right: S_0, D_{yy} and D_{yz}. Third row, left to right: FA, $< \mathbf{D} >$ and D_{zz}.

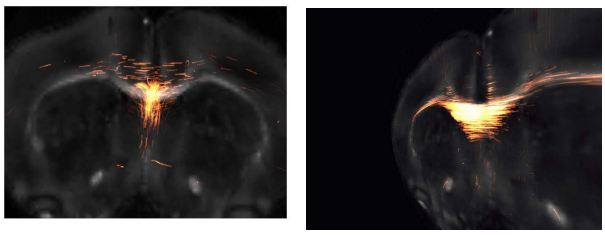


Fig. 3. Rat brain fiber tracts in and around the corpus callosum visualized using particle systems superimposed on the S_0 image. The particles are shown in bright orange. Left: Intermediate frame of an animation sequence depicting fiber growth; Right: the last frame of the same sequence from a different viewpoint. *[This is a color figure]*

5 Conclusions

We presented a novel constrained variational principle formulation for simultaneous smoothing and estimation of the positive definite diffusion tensor field from diffusion weighted images (DWI). To our knowledge, this is the first attempt at simultaneous smoothing and estimation of the positive definite diffusion tensor field from the raw data. We used the Cholesky decomposition to incorporate the positive definiteness constraint on the diffusion tensor to be estimated. The constrained variational principle formulation is transformed into a sequence of unconstrained problems using the augmented Lagrangian technique and solved numerically. Proof of the existence of a solution for each problem in the sequence of unconstrained problems is outlined.

Results of comparison between our method and a competing scheme [9] are shown for synthetic data under different situations involving use of linearized and nonlinear data acquisition models depicting the influence of the choice of the data acquisition model on the estimation. It was concluded that using the nonlinear data model yields better accuracy in comparison to the log-linearized model. Also, the superiority of our method in estimating tensor field over the chosen competing method was demonstrated for the synthetic data experiment.

Finally, fiber tract mapping of a normal rat brain were depicted using a particle system based visualization scheme. The estimated diffusion tensors are quite smooth without losing essential features when inspected visually. However, quantitative validation of the estimated fibers is essential and will be the focus of our future efforts.

Acknowledgments

Authors would like to thank Timothy McGraw in providing assistance on generating the fibers and Evren Ozarslan for providing us the rat brain DWI and for discussing the details of the physics of imaging.

References

1. L.Alvarez, P. L. Lions, and J. M. Morel, “Image selective smoothing and edge detection by nonlinear diffusion. ii,” *SIAM J. Numer. Anal.*, vol. 29, no. 3, pp. 845-866, June 1992.
2. P. J. Basser, J. Mattiello and D. Lebihan, “Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo,” *J. Magn. Reson.*, series B 103, pp. 247-254, 1994.
3. P. J. Basser and C. Pierpaoli, “Microstructural and Physiological Features of Tissue Elucidated by Quantitative-Diffusion-Tensor MRI,” *J. Magn. Reson.*, series B 111, pp. 209-219, 1996.
4. P. Blomgren, T.F. Chan and P. Mulet, “Extensions to Total Variation Denoising,”, Tech. Rep.97-42, UCLA, September 1997.
5. C. Chefd’hotel, D. Tschumperle’, Rachid Deriche, Olivier D. Faugeras, “Constrained Flows of Matrix-Valued Functions: Application to Diffusion Tensor Regularization,” *ECCV*, Vol. 1, pp. 251-265, 2002.
6. V. Caselles, J. M. Morel, G. Sapiro, and A. Tannenbaum, *IEEE TIP*, special issue on PDEs and geometry-driven diffusion in image processing and analysis, Vol 7, No. 3, 1998.

7. T.F. Chan, G. Golub, and P. Mulet, "A nonlinear primal-dual method for TV-based image restoration,"in *Proc. 12th Int. Conf. Analysis and Optimization of Systems: Images, Wavelets, and PDE's*, Paris, France, June 26-28, 1996, M. Berger et.al., Eds., no. 219, pp. 241-252.
8. T.E. Conturo, N.F. Lori, T.S. Cull, E. Akbudak, A.Z. Snyder, J.S. Shimony, R.C. McKinstry, H. Burton, and M.E. Raichle, "Tracking neuronal fiber pathways in the living human brain," in *Proc. Natl. Acad. Sci. USA 96*, 10422-10427 (1999)
9. O. Coulon, D.C. Alexander and S.R. Arridge, "A Regularization Scheme for Diffusion Tensor Magnetic Resonance Images," *IPMI*, 2001, pp. 92-105, Springer-Verlag .
10. L.C. Evans, "Partial Differential Equations," *Graduate Studies in Mathematics*, American Mathematical Society, 1997.
11. D.K. Jones, A. Simmons, S.C.R. Williams, and M.A. Horsfield, "Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI," *Magn. Reson. Med.*, 42, 37-41 (1999).
12. R. Kimmel, R. Malladi and N.A. Sochen, "Images as Embedded Maps and Minimal Surfaces: Movies, Color, Texture, and Volumetric Medical Images," *IJCV*, 39(2), pp. 111-129, 2000.
13. Bernard W.Lindgren *Statistical Theory*, Chapman & Hall/CRC, 1993.
14. J. Nocedal and S.J. Wright, *Num. Optimization*, Springer, 2000.
15. A. Pang and K. Smith, "Spray Rendering: Visualization Using Smart Particles," *IEEE Visualization 1993 Conference Proceedings*, 1993, pp. 283-290.
16. G.J.M. Parker, J.A. Schnabel, M.R. Symms, D.J. Werring, and G.J. Baker, "Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging," *Magn. Reson. Imag.* 11, 702-710, 2000.
17. P.Perona and J.Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE TPAMI*, vol. 12, no. 7, pp. 629–639, 1990.
18. Perona, P., "Orientation diffusions," *IEEE TIP*, vol.7, no.3, pp. 457-467, 1998.
19. C. Poupon, J.F. Mangin, C.A. Clark, V. Frouin, J. Regis, D.Le Bihan and I. Block, "Towards inference of human brain connectivity from MR diffusion tensor data,", *Med. Image Anal.*, vol. 5, pp. 1-15, 2001.
20. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259-268, 1992.
21. B. Tang, G. Sapiro and V. Caselles, "Diffusion of General Data on Non-Flat Manifolds via Harmonic Maps Theory: The Direction Diffusion Case," *IJCV*, 36(2), pp. 149-161, 2000.
22. D. Tschumperle and R. Deriche. "Regularization of orthonormal vector sets using coupled PDE's," *Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pp. 3-10, July 2001
23. B. C. Vemuri, Y. Chen, M.Rao, T.McGraw, Z. Wang and T.Mareci, "Fiber Tract Mapping from Diffusion Tensor MRI," *Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pp. 81-88, July 2001
24. J.Weickert, "A review of nonlinear diffusion filtering," *Scale-space theory in computer vision*, 1997, vol. 1252, of *Lecture Notes in Computer Science*, pp. 3–28, Springer-Verlag.
25. C.F. Westin, S.E. Maier, H. Mamata A. Nabavi, F.A. Jolesz and R. Kikinis, "Processing and visualization for diffusion tensor MRI," *Med. Image Anal.*, vol. 6, pp. 93-108, 2002.