

Non-stationary Multi-output Gaussian Processes for Enhancing Resolution over Diffusion Tensor Fields

Jhon F. Cuellar-Fierro¹(✉), Hernán Darío Vargas-Cardona¹,
Mauricio A. Álvarez², Andrés M. Álvarez¹, and Álvaro A. Orozco¹

¹ Automatic Research Group, Universidad Tecnológica de Pereira, Pereira, Colombia
{jfcuellar, hernan.vargas, andres.alvarez1, aaog}@utp.edu.co

² University of Sheffield, Sheffield, UK
mauricio.alvarez@sheffield.ac.uk

Abstract. Diffusion magnetic resonance imaging (dMRI) is an advanced technique derived from magnetic resonance imaging (MRI) that allows the study of internal structures in biological tissue. Due to acquisition protocols and hardware limitations of the equipment employed to obtain the data, the spatial resolution of the images is often low. This inherent lack in dMRI is a considerable difficulty because clinical applications are affected. The scientific community has proposed several methodologies for enhancing the spatial resolution of dMRI data, based on interpolation of diffusion tensors fields. However, most of the methods have considerable drawbacks when they interpolate strong transitions, such as crossing fibers. Also, relevant clinical information from tensor fields is modified when interpolation is performed. In this work, we propose a probabilistic methodology for interpolation of diffusion tensors fields using multi-output Gaussian processes with non-stationary kernel function. First, each tensor is decomposed in shape and orientation features. Then, the model interpolates the features jointly. Results show that proposed approach outperforms state-of-the-art methods regarding resolution enhancement accuracy on synthetic and real data, when we evaluate interpolation quality with Frobenius and Riemann metrics. Also, the proposed method demonstrates an adequate characterization of both stationary and non-stationary fields, contrary to previous approaches where performance is seriously reduced when complex fields are interpolated.

1 Introduction

Since the appearance of Magnetic Resonance Imaging (MRI) several years ago, many clinical applications have been developed. As well as a variety of methodologies for improving diagnosis of various neurological diseases, e.g., Parkinson, Epilepsy, [1]. Recently, the diffusion MRI (dMRI) has emerged as a non-invasive method that allows exploring, visualizing, and evaluating qualitatively and quantitatively biological structures, such as white matter tracts, cortical gray substance, cardiac fibers, among others [2]. dMRI describes the diffusion of water

particles through a 2^{nd} order tensor $\mathbf{D} \in \mathbb{R}^{3 \times 3}$, which fully represents the particle mobility along each spatial direction, yielding anisotropic diffusions [3]. Therefore, the diffusion tensor of each voxel in a given dMRI is represented by a 3×3 symmetric and positive definite matrix, being necessary at least six independent measurements along different directions.

However, dMRI has a considerable difficulty with the spatial resolution of acquired images. This problem is due to clinical acquisition protocols, and technological limitations of MRI scanners [4]. Spatial resolution of dMRI is commonly in a range of 1 to 2 mm³ for each voxel. Nonetheless, in medical imaging it is often desired to obtain information from smaller structures, leading to a lack of precision in some clinical applications [5]. Different methodologies for interpolation of diffusion tensor fields have been proposed for enhancing dMRI resolution. Diffusion tensors have to satisfy some restrictions. For example, the determinant of matrices must change monotonously to avoid the swelling effect [5], and they must be positive definite (PD). The first attempt for improving resolution of dMRI is based on direct interpolation, where each component of the tensor is interpolated independently in a Euclidean space [6]. However, this technique generates swelling effect [6]. To preserve some constraints of diffusion tensors (i.e., PD tensors), parametric approaches have been developed using Cholesky factorization [7]. Nevertheless, the fractional anisotropy (FA) is modified. Other proposed approaches [8] simultaneously seek to eliminate the swelling effect and preserve the symmetry and PD property of diffusion tensors by interpolation in Riemannian spaces. However, their computational cost is high. To overcome computational cost issue and PD constraint, the Log-Euclidean interpolation [6] applies a logarithmic transformation to the matrices before operating on them, but its performance is reduced in heterogeneous fields. An alternative study called Feature-based interpolation (FbI), proposed by [5], developed an approach for interpolation of diffusion tensors by decomposing the tensors into eigenvalues (direction) and Euler angles (orientation). Then, each feature is interpolated linearly. The previously mentioned methods [5–8], have difficulties to describe non-stationary fields: strong transitions and abrupt changes among tensors such as crossing fibers.

In this work, we present a methodology for interpolation of diffusion tensor fields based on multi-output Gaussian processes with a non-stationary kernel function (NmoGp). First, we decompose each tensor of the field in three eigenvalues and three Euler angles, following the method proposed in [5]. Thus, each tensor is represented by six features indexed in an independent variable that represents spatial coordinates (x, y, z) . The main goal is to describe non-stationary diffusion tensor fields allowing an accurate interpolation of complex and noisy dMRI data. To do this, we use an expressive kernel to construct the covariance matrix of the multi-output Gaussian process. The expressive kernel is a convex combination of several kernels or the same kernel with different hyperparameters. The methodology introduced is an extension of [9] that we call moGp. In this case, we are interested in a more robust description of complex fields (i.e. crossing fibers) using a non-stationary kernel function. We compare our

approach (NmoGp) against the stationary moGp [9] and the FbI proposed by [5], evaluating Frobenius and Riemann distances.

2 Materials and Methods

2.1 dMRI and Tensor Fields

The dMRI is a measurement of diffusion of water particles inside biological tissues. This diffusion is fully described using a 2^{nd} order tensor $\mathbf{D} \in \mathbb{R}^{3 \times 3}$, which defines shape, orientation and direction [3]. The diffusion tensor is expressed as a 3×3 symmetric positive definite matrix whose elements $D_{ij} = D_{ji}$ with $i, j \in \{x, y, z\}$. It is necessary six independent gradient diffusion measures ($S_k \in \mathbb{R}^+$, $k = \{1, \dots, 6\}$) and a baseline image without gradient field ($S_0 \in \mathbb{R}^+$) for estimating the coefficients of a tensor (D_{ij}) from Stejskal-Tanner equation [3]: $S_k(\mathbf{x}) = S_0(\mathbf{x})e^{-b\mathbf{g}_k^\top D(\mathbf{x})\mathbf{g}_k}$, where, S_k is k^{th} dMRI, $\mathbf{g}_k \in \mathbb{R}^3$ is a factor that controls strength and timing of gradients.

2.2 Feature-Based Interpolation Framework

A scheme for tensorial interpolation that keeps restrictions of a diffusion tensor was proposed by [5]. It consists of decomposing a diffusion tensor in six features: three eigenvalues (direction) and three Euler Angles (orientation). Each tensor of the field is decomposed as $\mathbf{D} = \mathbf{E}\Lambda\mathbf{E}^\top$, where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^+$ is a diagonal matrix of eigenvalues and \mathbf{E} is a matrix whose columns holds the eigenvector of the tensor \mathbf{D} as follows:

$$\mathbf{E} = \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix}. \quad (1)$$

For preserving the PD property, we apply a logarithmic transformation to eigenvalues as suggested in [5]. A tensor can be represented as a feature vector $\mathbf{y} \in \mathbb{R}^6$ indexed in positions coordinates, $\mathbf{x} = [x, y, z]^\top$, yielding: $\mathbf{y}(\mathbf{x}) = [\ln \lambda_1(\mathbf{x}), \ln \lambda_2(\mathbf{x}), \ln \lambda_3(\mathbf{x}), \alpha(\mathbf{x}), \beta(\mathbf{x}), \gamma(\mathbf{x})]^\top$, where Euler angles are given by: $\alpha = \text{artan2}(v_{12}, v_{11})$, $\beta = \text{artan2}(-v_{13}, \sqrt{v_{11}^2 + v_{12}^2})$, $\gamma = \text{artan2}(v_{23}, v_{33})$ and $\text{artan2}(a, b)$ is the four-quadrant arctangent of the real arguments a and b .

2.3 Multi-output Gaussian Processes and Non-stationary Kernel

The model for multi-output is defined as a collection of random variables. Such that any finite of them follows a joint Gaussian distribution. The random variables are associated with a set of different processes $\{f_d\}_{d=1}^{M=6}$, evaluated at different values of \mathbf{x} [10]. Therefore, it is assumed the vector-value function \mathbf{f} as a Gaussian process. Mathematically is $\mathbf{f} \sim \mathcal{GP}(\mathbf{m}(\mathbf{X}), \mathbf{K}(\mathbf{X}, \mathbf{X}))$ where $\mathbf{m}(\mathbf{X})$ is a M -dimensional vector whose components are the mean functions $\{m_d(\mathbf{x})\}_{d=1}^{M=6}$ of each output, and $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is a $NM \times NM$ positive defined matrix, with entries,

$(\mathbf{K}(x_i, x_j))_{d,d'}$, for $i, j = 1, \dots, N$ and $d, d' = 1, \dots, M$, (being N and M the number of training samples and outputs). The predictive distribution for a new input vector x_* is a Gaussian distribution [11] and it is given by (2):

$$p(\mathbf{f}(x_*)|\mathbf{S}, \mathbf{f}, x_*, \phi) = \mathcal{N}(\mathbf{f}_*(x_*), \mathbf{K}_*(x_*, x_*)) \quad (2)$$

with, $\mathbf{f}_*(x_*) = \mathbf{K}_{x_*}^\top (\mathbf{K}(\mathbf{X}, \mathbf{X}) + \mathbf{\Sigma})^{-1} \bar{\mathbf{y}}$, and $\mathbf{K}_*(x_*, x_*) = \mathbf{K}(x_*, x_*) - \mathbf{K}_{x_*}^\top (\mathbf{K}(\mathbf{X}, \mathbf{X}) + \mathbf{\Sigma})^{-1} \mathbf{K}_{x_*}$, where $\mathbf{K}_{x_*} \in \mathbf{R}^{M \times NM}$ has entries $(\mathbf{K}(x_*, x_j))_{d,d'}$ $i = 1, \dots, N$ and $d, d' = 1, \dots, M$, $\mathbf{\Sigma}$ is a diagonal matrix whose elements are noise of the observation, $\bar{\mathbf{y}}$ is an NM vector obtained concatenating the output vectors, \mathbf{S} is the training data, and ϕ represents the set of hyperparameters of the covariance function and the variances of noise for each output $\{\sigma_d^2\}_{d=1}^{M=6}$. There are several possible choices of covariance for multi-output problem, in this work we consider the linear model of coregionalization (LMC) [10]. Usually, the kernel employed in moGp is one of the common stationary functions available in the literature (RBF, Matern, rational quadratic, among others). On the other hand, we define a non-stationary kernel as an expressive mixture according to the model proposed by [12]:

$$k(x, x') = \sum_{i=1}^r \sigma(w_i(x)) k_i(x, x') \sigma(w_i(x')), \quad (3)$$

where $w_i(x) : \mathbb{R}^P \rightarrow \mathbb{R}^1$ is the weighting function, with P the dimensional input. The expressiveness of this function determines how many changes can occur in the data. $w_i(x) = \sum_{j=1}^v a_j \cos(\omega_j x + b_j)$. $\sigma(z) : \mathbb{R}^1 \rightarrow [0, 1]$, is the warping function, that is computed as a convex combination over the weighting function $\sigma(w_i(x)) = \exp(w_i(x)) / \sum_{i=1}^r \exp(w_i(x))$, $\sum_{i=1}^r \sigma(w_i(x)) = 1$ inducing a partial discretization over latent functions. This function produces non-stationarity, since it depends of the input variable x , and $k_i(x, x')$ can be any stationary kernel. It is expected that latent functions have different kernel structures or a same form with different hyperparameters.

2.4 Experimental Setup and Datasets

First, we test the proposed model in a simulation of crossing fibers. This dataset is obtained from FanDTasia Toolbox [13]. Also, we evaluated real dMRI data of the head from a healthy subject. We employ 25 gradient directions with a value b equal to $1000\text{S}/\text{mm}^2$. Both datasets are splitted in a 50% for training and 50% for validation. Original datasets are the gold standard fields. We compare our model with methodologies proposed by [5] and [9]. Finally, to quantitatively evaluate the models, we use two distance metrics defined over matrices: the Frobenius norm (Frob) and Riemann distance (Riem) [6], given by (4) and (5).

$$\text{Frob}(\mathbf{D}_1, \mathbf{D}_2) = \sqrt{\text{trace} \left[(\mathbf{D}_1 - \mathbf{D}_2)^\top (\mathbf{D}_1 - \mathbf{D}_2) \right]}, \quad (4)$$

$$\text{Riem}(\mathbf{D}_1, \mathbf{D}_2) = \sqrt{\text{trace} \left[\log(\mathbf{D}_1^{-1/2} \mathbf{D}_2 \mathbf{D}_1^{-1/2})^\top \log(\mathbf{D}_1^{-1/2} \mathbf{D}_2 \mathbf{D}_1^{-1/2}) \right]}, \quad (5)$$

where \mathbf{D}_1 and \mathbf{D}_2 are the interpolated and the gold standard tensor, respectively. The error metrics are computed in all the test data, and we report the mean, standard deviation and confidence interval for each model.

3 Experimental Results and Discussion

3.1 Crossing Fibers

Crossing fibers data is one of most difficult field to interpolate, because transition among tensors is very abrupt. We test the NmoGp over a 2D crossing fibers field of 22×22 . We use five latent functions ($r = 5$) and we employ a rational quadratic kernel. The Fig. 1(a) shows the crossing fibers field, Fig. 1(b) is the downsampled tensor field (used for training), and Fig. 1(c) illustrates the validation tensors. Also, we show Riemann error maps in Figs. 1(d), (e) and (f). If we observe in detail, error values are very low over smooth regions of the tensor field for all methods (blue zones in error maps of Fig. 1). However, when we evaluate strong transition regions, interpolation is not straightforward, and the error is higher in this areas. Specifically, over this critic region of the field, our model achieves more precision than comparison methods. We can explain this because the non-stationary kernel can capture the dynamic of the tensor field. While it is true that there are some errors when we interpolate new tensors in the crossing fibers region, we can say that NmoGp outperforms to moGp and the FbI method, when interpolation is challenging due to strong changes in neighboring tensors.

Table 1. Frob and Riem distances for crossing fibers of NmoGp, moGp, and FbI methods

| Model | Riem | Confidence interval | Frob (10e-4) | Confidence interval |
|-------|-------------------|---------------------|---------------------|---------------------|
| NmoGp | 0.191 ± 0.277 | [0.156, 0.227] | 17.342 ± 25.191 | [1.410, 2.058] |
| moGp | 0.213 ± 0.283 | [0.176, 0.249] | 19.589 ± 26.136 | [1.622, 2.294] |
| FbI | 0.266 ± 0.362 | [0.220, 0.313] | 24.228 ± 33.375 | [1.993, 2.851] |

Also, according to results in Table 1, the FbI model obtained the higher error (Frobenius and Riemann). FbI model has a considerable lack because it does not consider correlation among the six features extracted from tensor decomposition. This approach only interpolates linearly and separately each feature. On the other hand, moGp and NmoGp interpolate the six features simultaneously. The idea is to share a correlation among features. For this reason, there is an

additional information allowing a better estimation of new data. The main difference between moGp and NmoGP is the type of kernel for constructing the covariance matrix inside the model. Thereby, moGp works with a single kernel (stationary) while our NmoGP is based on a non-stationary kernel. We see in Table 1, that global error of the NmoGP approach is lower than the moGp and FbI method.

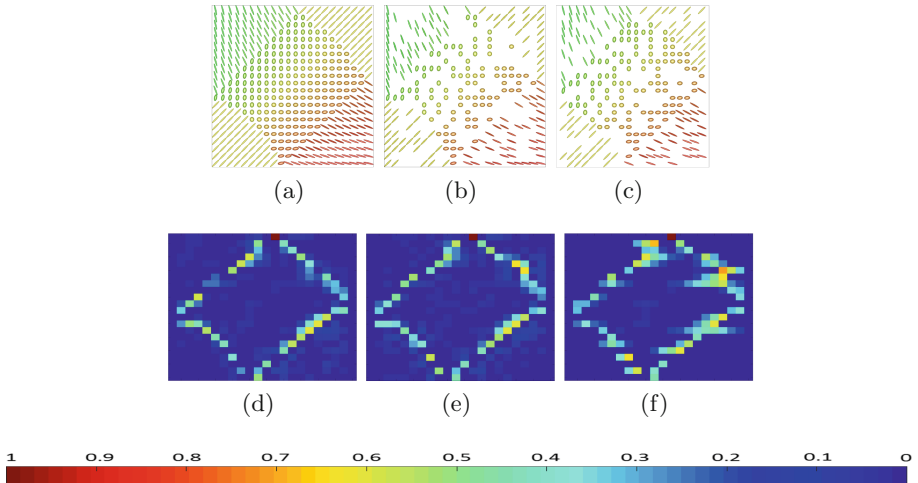


Fig. 1. Interpolation over crossing fiber fields, (a) gold standard, (b) training field, (c) test field. (d), (e) and (f) Riemann norm error maps, for the NmoGp, moGp and FbI model, respectively.

3.2 Real DT Field

We test the methods over a 2D diffusion tensor field obtained from real dMRI. The data corresponds to an axial slice with 40×40 tensors of the head. Figures 2(a), (b) and (c) correspond to the gold standard, training and test data respectively. Figures 2(d), (e) and (f) show the Riemann error maps. Again, Table 2 show the error distance for all tested methods and the confidence interval. A real dMRI tensor field is noisy and very heterogeneous. Therefore, simple interpolation methods fail to achieve a good accuracy. Also, probabilistic methods such as moGp and NmoGP have the robustness property. Again, NmoGP improves to the comparison methods according to outcomes of Table 2. This result is very relevant, because it confirms that proposed method can interpolate tensorial data with good accuracy, no matter the type of dataset. Finally, we can establish that insertion of a non-stationary kernel in multi-output Gaussian process increases its performance significantly.

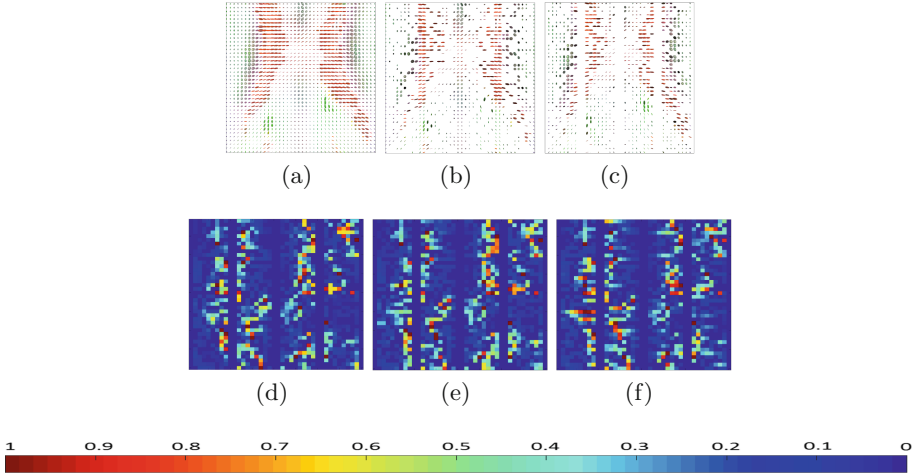


Fig. 2. Interpolation over a 2D real crossing fiber field, (a) gold standard, (b) training field, (c) test field. (d), (e) and (f) Riem norm error maps for the NmoGp, moGp and FbI model, respectively.

Table 2. Frob and Riem distances for real data of NmoGp, moGp, and FbI method

| Model | Riem | Confidence interval | Frob | Confidence interval |
|-------|--------------------|---------------------|-------------------|---------------------|
| NmoGp | 11.222 ± 1.041 | [1.051, 1.193] | 0.725 ± 0.553 | [0.686, 0.764] |
| moGp | 11.458 ± 0.958 | [1.078, 1.213] | 0.745 ± 0.497 | [0.709, 0.780] |
| FbI | 11.569 ± 1.044 | [1.085, 1.227] | 0.756 ± 0.504 | [0.720, 0.791] |

4 Conclusions and Future Work

In this work, we presented a probabilistic methodology for interpolation of diffusion tensor fields. The model decomposes the tensors in six features describing the main properties of a diffusion tensor (direction and orientation). We index the extracted features in spatial coordinates to interpolate them using a moGp with a non-stationary kernel function. The structure of the kernel function employed in this method combines several kernels with different properties. The purpose is to characterize complex fields such as crossing fibers. In this context, the non-stationary kernel allows differentiating between strong and uniform transitions. In this way, the interpolation of new tensors is more accurate in comparison to ordinary approaches. We tested NmoGp against state-of-the-art methods in two different datasets: simulation of crossing fibers, and a dMRI segment. Outcomes proved that proposed model outperforms to the comparison methods evaluating accuracy with Frobenius and Riemann distances. Although, our proposed approach is accurate and robust for interpolation of complex tensor fields, there

is an important issue: initialization of parameters and hyperparameters is not straightforward. Currently, we use a cross-validation procedure. As future work, we would like to extend non-stationary kernel functions to more complex models such as generalized Wishart processes and tractography procedures.

Acknowledgments. This research is developed under the project “Desarrollo de un sistema de soporte clínico basado en el procesamiento estocástico para mejorar la resolución espacial de la resonancia magnética estructural y de difusión con aplicación al procedimiento de la ablación de tumores” financed by COLCIENCIAS with code 111074455860 under the program: *744 Convocatoria para proyectos de ciencia, tecnología e innovación en salud 2016*. H.D. Vargas is funded by Colciencias under the program: *Formación de alto nivel para la ciencia, la tecnología y la innovación - Convocatoria 617 de 2013* with code 111065740687. We thank to the program of master in electrical engineering of the Universidad Tecnológica de Pereira.

References

1. Sasiadek, M.J., Szewczyk, P., Bładowska, J.: Application of diffusion tensor imaging (DTI) in pathological changes of the spinal cord. *Med. Sci. Monit.* **18**(6), RA73–RA79 (2012)
2. Del Cura, J.L., Pedraza, S.: *Radiología Esencial*. Toma 2 (2010)
3. Le Bihan, D., Mangin, J.-F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabriat, H.: Diffusion tensor imaging: concepts and applications. *J. Magn. Reson. Imaging* **13**(4), 534–546 (2001)
4. Gupta, V., Ayache, N., Pennec, X.: Improving DTI resolution from a single clinical acquisition: a statistical approach using spatial prior. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *MICCAI 2013*. LNCS, vol. 8151, pp. 477–484. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40760-4_60
5. Yang, F., Zhu, Y.-M., Magnin, I.E., Luo, J.-H., Croisille, P., Kingsley, P.B.: Feature-based interpolation of diffusion tensor fields and application to human cardiac DT-MRI. *Med. Image Anal.* **16**(2), 459–481 (2012)
6. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.* **56**(2), 411–421 (2006)
7. Wang, Z., Vemuri, B.C., Chen, Y., Mareci, T.H.: A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. *IEEE Trans Med. Imag.* **23**(8), 930–939 (2004)
8. Batchelor, P.G., Moakher, M., Atkinson, D., Calamante, F., Connelly, A.: A rigorous framework for diffusion tensor calculus. *Magn. Reson. Med.* **53**(1), 221–225 (2005)
9. Cardona, H.D.V., Orozco, Á.A., Álvarez, M.A.: Multi-output Gaussian processes for enhancing resolution of diffusion tensor fields. In: 2016 IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), pp. 1111–1114. IEEE (2016)
10. Alvarez, M.A., Rosasco, L., Lawrence, N.D., et al.: Kernels for vector-valued functions: a review. *Found. Trends® Mach. Learn.* **4**(3), 195–266 (2012)
11. Rasmussen, C.E.: *Gaussian Processes for Machine Learning*. MIT Press, Cambridge (2006)

12. Herlands, W., Wilson, A., Nickisch, H., Flaxman, S., Neill, D., Van Panhuis, W., Xing, E.: Scalable Gaussian processes for characterizing multidimensional change surfaces. arXiv preprint [arXiv:1511.04408](https://arxiv.org/abs/1511.04408) (2015)
13. Barmpoutis, A., Vemuri, B.C.: A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints. In: Biomedical Imaging From Nano to Macro, pp. 1385–1388. IEEE (2010)