Higher-Order Tensors in Diffusion Imaging

Thomas Schultz, Andrea Fuster, Aurobrata Ghosh, Rachid Deriche, Luc Florack, Lim Lek-Heng

Higher-Order Tensors in Diffusion Imaging

Thomas Schultz, Andrea Fuster*, Aurobrata Ghosh*, Rachid Deriche, Luc Florack, Lek-Heng Lim

Abstract Diffusion imaging is a noninvasive tool for probing the microstructure of fibrous nerve and muscle tissue. Higher-order tensors provide a powerful mathematical language to model and analyze the large and complex data that is generated by its modern variants such as High Angular Resolution Diffusion Imaging (HARDI) or Diffusional Kurtosis Imaging. This survey gives a careful introduction to the foundations of higher-order tensor algebra, and explains how some concepts from linear algebra generalize to the higher-order case. From the application side, it reviews a variety of distinct higher-order tensor models that arise in the context of diffusion imaging, such as higher-order diffusion tensors, q-ball or fiber Orientation Distribution Functions (ODFs), and fourth-order covariance and kurtosis tensors. By bridging the gap between mathematical foundations and application, it provides an introduction that is suitable for practitioners and applied mathematicians alike, and propels the field by stimulating further exchange between the two.

1 Introduction

In biological tissues such as nerve fiber bundles and muscles, the spontaneous heat motion of water molecules is restricted by obstacles in the fibrous microstructure.

Thomas Schultz
University of Bonn, Germany. e-mail: schultz@cs.uni-bonn.de

Andrea Fuster* and Luc Florack
Eindhoven University of Technology, The Netherlands. e-mail: {a.fuster,l.m.j.florack}@tue.nl

Aurobrata Ghosh* and Rachid Deriche
INRIA Sophia Antipolis-Méditerranée, France. e-mail: {aurobrata.ghosh,rachid.deriche}@inria.fr

Lek-Heng Lim
University of Chicago, USA. e-mail: lekheng@uchicago.edu

* These authors have contributed equally.

Diffusion Imaging [70] uses the principles of Magnetic Resonance Imaging (MRI) to non-invasively measure properties of this motion, which is also known as self-diffusion. When applied to the human brain, this provides unique insights about brain connectivity, which makes diffusion MRI one of the key technologies in an ongoing large-scale scientific effort to map the human brain connectome [33]. Consequently, it is a timely and important topic of research to create mathematical models that infer biologically meaningful parameters from such data.

Higher-order tensors have been used in applications ranging from psychometrics [64] and chemometrics [104] to signal processing [103], computer vision [111], and neuroscience [85]. They also provide adequate models for a number of quantities that occur in the context of diffusion imaging. Many practitioners view higher-order tensors as a generalization of matrices to multi-way arrays. However, tensors can also be studied in an invariant, coordinate-free notation. Tensor decompositions are an active and challenging topic in applied mathematics, since fundamental concepts from linear algebra, such as the singular value decomposition, do not have unique generalization to higher order, and most generalizations are hard to compute.

It is a goal of our survey to stimulate an active exchange between mathematicians, who are studying tensor decompositions and the geometry of tensors, and computer scientists and MR physicists, who are interested in using tensors as mathematical tools in the context of diffusion MRI. Therefore, unlike previous surveys [90, 39], Section 2 provides a broad overview of all physical quantities that have been modeled with higher-order tensors in the context of diffusion MRI. On the other hand, our introduction to the higher-order tensor formalism in Section 3 differs from existing discussions [66, 73] by focusing on aspects relevant to this specific application.

Relevant literature is spread over journals in applied mathematics, MR physics, neuroimaging, and computer science. Drawing on all these fields, Section 4 presents the current state of the art on fitting higher-order tensor models to the measured data, and Section 5 discusses operations performed on the tensors for further analysis. Among others, this includes computation of scalar invariants (Section 5.1), maximum detection (Section 5.3), and tensor decompositions (Section 5.4).

2 Overview of Higher-Order Tensor Models in dMRI

Different physical quantities that can be measured by or inferred from diffusion MRI have been modeled with higher-order tensors. The resulting tensors not only differ in their interpretation, but also in dimension, order, and symmetry.

Diffusion imaging inserts magnetic field gradients into the MR sequence which sensitize the measurement to molecular motion along the gradient direction [70]. Compared to an image without diffusion weighting, this leads to an attenuation of signal strength. The standard diffusion tensor model [18] assumes that the diffusion-weighted MR signal in direction $\mathbf{u}$ is given by a monoexponential attenuation of the unweighted signal S_0, depending on the diffusion weighting b and a directionally dependent apparent diffusion coefficient, modeled by a diffusion tensor $\mathbf{D}$:

$$S(\mathbf{u}) = S_0 e^{-b\mathbf{u}^T\mathbf{D}\mathbf{u}} \tag{1}$$

Estimating the six unique coefficients of $\mathbf{D}$ requires measurements in at least six different gradient directions. Typical parameter values are $b \in [700, 1000]\, \text{mm}^2/\text{s}$, and a spatial resolution of around $2 \times 2 \times 2\, \text{mm}^3$. When studying the human brain, this corresponds to a subdivision into around 10^5 volume elements (voxels); a separate diffusion tensor $\mathbf{D}$ is computed for each of them.

Since nerve fibers are on the micrometer scale and therefore far below image resolution, their complex organization often leads to apparent diffusivities $D(\mathbf{u})$ that are poorly approximated by a quadratic function. For these cases, Eq. (1) has been generalized to use higher-order polynomials. As it will be explained in Section 3.3, this corresponds to a higher-order diffusion tensor $\mathscr{D}$ [86]:

$$S(\mathbf{u}) = S_0 e^{-bD(\mathbf{u})} \quad \text{with} \quad D(\mathbf{u}) = \mathscr{D} \cdot^k \mathbf{u} \tag{2}$$

Such High Angular Resolution Diffusion Imaging (HARDI) models require a larger number of 30–100 gradient directions, and larger $b \in [1000, 3000]\, \text{mm}^2/\text{s}$.

One goal in diffusion imaging is to estimate the dominant nerve fiber directions within each voxel. When there is only one such direction, the principal eigenvector of the diffusion tensor $\mathbf{D}$ is aligned with it. However, a mixture of multiple fiber directions is not easily resolved with the higher-order diffusion tensor $\mathscr{D}$. For this purpose, it is easier to consider the diffusion propagator $P(\mathbf{x})$, the probability density of a molecular displacement along vector $\mathbf{x}$ within the diffusion time. Under certain assumptions, $P(\mathbf{x})$ can be computed from $\mathscr{D}$; this will be the topic of Section 5.2.

Writing the diffusion propagator $P(\mathbf{x})$ in spherical coordinates and integrating over the radius results in the diffusion orientation distribution function $\psi(\mathbf{u})$, whose maxima approximate the main nerve fiber directions. The q-ball model has been introduced as an approximative way of computing $\psi(\mathbf{u})$ [110]. Even though its exact interpretation has been disputed [17], q-ball maxima indicate approximate fiber directions, and q-balls are sometimes expressed in a tensor basis [55, 7], making it relevant to compute the maxima of homogeneous forms (cf. Section 5.3).

When measuring at different b values, it is common to observe that the true signal attenuation is not monoexponential, as assumed by Eqs. (1) and (2). This indicates that the diffusion propagator $P(\mathbf{x})$ is non-Gaussian. Accounting for all higher-order moments of P leads to a different generalization of Eq. (1) [78, 77],

$$S(\mathscr{B}) = S_0 \, e^{\sum_{k=2}^{\infty} j^k \langle \mathscr{D}^{(k)}, \mathscr{B}^{(k)} \rangle}, \tag{3}$$

where j is the imaginary unit, $\mathscr{D}^{(k)}$ is a series of diffusion tensors with increasing order k, and the diffusion-weighted signal $S(\mathscr{B})$ is a function of a series of tensors $\mathscr{B}^{(k)}$, which combine information about the direction and strength of the diffusion weighting. $\langle \mathscr{D}^{(k)}, \mathscr{B}^{(k)} \rangle$ denotes the scalar product of the two tensors.

In contrast to Eq. (2), which uses a single higher-order tensor $\mathscr{D}$ that contains all the information that would be present in lower-order approximations, each element in the series of tensors $\mathscr{D}^{(k)}$ in Eq. (3) contains non-redundant information that is in-

dependent from all other orders k. This additional information needs to be acquired by sampling multiple b values in several gradient directions [79].

The tensors in Eq. (3) are three-dimensional, and symmetric under all index permutations. The odd orders k in Eq. (3) carry information about asymmetries in the diffusion propagator, i.e., $P(-\mathbf{x}) \neq P(\mathbf{x})$. However, that information resides in the phase of the complex-valued MR signal. At the technical state of the art, signal phase in diffusion MRI is so heavily corrupted by measurement noise and artifacts that it is not informative. Therefore, practical implementations of this generalization are limited to estimating even-order tensors from the signal magnitude [80].

Diffusional Kurtosis Imaging augments the second-order diffusion tensor $\mathbf{D}$ in Eq. (1) with a fourth-order kurtosis tensor $\mathscr{W}$ [61],

$$S(\mathbf{u},b) = S_0 \mathrm{e}^{-b\mathbf{u}^{\mathrm{T}}\mathbf{Du}+\frac{1}{6}b^2\left(\frac{1}{3}\mathrm{tr}(\mathbf{D})\right)^2 \mathscr{W}.^4\mathbf{u}}, \tag{4}$$

where tr indicates matrix trace. Computing the parameters in Eq. (4) requires measurements at multiple b values, but no signal phase. They capture the same information present in the second and fourth moments of $P(\mathbf{x})$, but allow for simpler computation of the apparent diffusional kurtosis K_{app} in direction $\mathbf{u}$:

$$K_{\mathrm{app}}(\mathbf{u}) = \frac{\left(\frac{1}{3}\mathrm{tr}(\mathbf{D})\right)^2}{\left(\mathbf{u}^{\mathrm{T}}\mathbf{Du}\right)^2} \mathscr{W}.^4\mathbf{u} \tag{5}$$

For Gaussian diffusion, $K_{\mathrm{app}} = 0$. Negative kurtosis is expected from diffusion restricted by spherical pores, and positive kurtosis can indicate presence of heterogeneous diffusion compartments [61].

Fourth-order covariance tensors Σ occur in statistical models of second-order diffusion tensors [19]. Even though they are three-dimensional in each mode, they only possess partial symmetries ($\Sigma_{ijkl} = \Sigma_{klij}$; $\Sigma_{ijkl} = \Sigma_{jikl}$; $\Sigma_{ijkl} = \Sigma_{ijlk}$) [20].

If we assume that all nerve fiber bundles within a voxel have approximately the same diffusion characteristics, the MR signal is given by the convolution of a fiber orientation density function (fODF) with a kernel describing the single fiber response [108]. Unlike the diffusion ODF, values of the fODF $F(\mathbf{u})$ are interpreted as the fraction of fibers aligned with direction $\mathbf{u}$. $F(\mathbf{u})$ can be obtained by spherical deconvolution and a variant of that technique, which will be explained in Section 4.3, allows for further analysis of the fODF via tensor decomposition [101].

3 Mathematical Background

We include a basic introduction to tensors and tensor fields. In a nutshell, a tensor of order p or p-tensor is a multilinear functional on p vector spaces $T : \mathbb{V} \times \mathbb{V} \times \cdots \times \mathbb{V} \to \mathbb{R}$, and can be represented in coordinates as a p-dimensional matrix $A \in \mathbb{R}^{n \times n \times \cdots \times n}$, $n = \dim(\mathbb{V})$, if one chooses a basis on $\mathbb{V}$. A tensor field is a tensor-valued function on a manifold. We refer readers who are interested in further properties

of tensors and hypermatrices to [73] for an elementary treatment. Mathematically sophisticated readers may consult [66] for a much more in-depth treatment.

3.1 Basic Definitions

Let us first define our basic mathematical objects: (i) tensors, and (ii) tensor fields. Let $\mathbb{V}$ be a vector space over $\mathbb{R}$. An order-p *tensor* is a *multilinear functional*

$$f : \underbrace{\mathbb{V} \times \mathbb{V} \times \cdots \times \mathbb{V}}_{p \text{ times}} \to \mathbb{R}.$$

Multilinear means that if all arguments are kept constant but one, then f is linear in that varying argument, i.e.

$$f(\mathbf{u}_1, \ldots, \alpha\mathbf{v}_i + \beta\mathbf{w}_i, \ldots, \mathbf{u}_p) = \alpha f(\mathbf{u}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{u}_p) + \beta f(\mathbf{u}_1, \ldots, \mathbf{w}_i, \ldots, \mathbf{u}_p), \tag{6}$$

for every $i = 1, \ldots, p$, $\alpha, \beta \in \mathbb{R}$ and $\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i \in \mathbb{V}$. The set of all p-tensors is called the p-fold *tensor product* of the vector space $\mathbb{V}$ and denoted

$$\mathbb{V}^{\otimes p} = \underbrace{\mathbb{V} \otimes \mathbb{V} \otimes \cdots \otimes \mathbb{V}}_{p \text{ times}}.$$

We ignore the distinction between covariant, contravariant and mixed tensors, since it is less relevant when working with coordinate representations in an orthonormal basis, as will be the case in this survey. An abstract approach towards tensors is now standard in any basic graduate courses in algebra [57, 68] or even mathematical methods courses for physicists [38]. However, such courses focus almost exclusively on properties of an entire space of tensors [117] as opposed to properties of an individual tensor, i.e. a specific element from such a tensor space. Properties of an individual tensor such as rank, norm, eigenvalues, decompositions, are of great relevance to us and will be discussed after we introduce tensor fields.

We will be informal in our treatment of tensor fields to make it more easily accessible. Readers who wish to see a rigorous definition would have no shortage of standard references [24, 67, 113] to consult. Let M be a topological manifold which we may later endow with additional structures (differential, Riemannian, Finsler, etc). A tensor field is, roughly speaking, a tensor-valued function $F : M \to \mathbb{V}^{\otimes p}$ or alternatively, a function of the form

$$F : M \times \underbrace{\mathbb{V} \times \mathbb{V} \times \cdots \times \mathbb{V}}_{p \text{ times}} \to \mathbb{R} \tag{7}$$

with the property that for every point $\mathbf{x} \in M$,

$$F(\mathbf{x}; \cdot, \cdot, \ldots, \cdot) : \mathbb{V} \times \mathbb{V} \times \cdots \times \mathbb{V} \to \mathbb{R}$$

is a multilinear functional, i.e., $F(\mathbf{x}; \mathbf{u}_1, \ldots, \mathbf{u}_p)$ is multilinear in the last p arguments for every fixed $\mathbf{x} \in M$. If we want F to have additional properties like continuity or differentiability, this definition is only good locally, i.e., every $\mathbf{x}_0 \in M$ has a neighborhood $U_{\mathbf{x}_0} \subseteq M$ such that

$$F : U_{\mathbf{x}_0} \times \mathbb{V} \times \mathbb{V} \times \cdots \times \mathbb{V} \to \mathbb{R}$$

is multilinear for every $\mathbf{x} \in U_{\mathbf{x}_0}$. By far the most common choice for $\mathbb{V}$ is $T_{\mathbf{x}}(M)$, the tangent space at $\mathbf{x}$, i.e., the vector space $\mathbb{V}_i$ changes with each $\mathbf{x}$ and we really have a multilinear function

$$F(\mathbf{x}; \cdot, \cdot, \ldots, \cdot) : T_{\mathbf{x}}(M) \times \cdots \times T_{\mathbf{x}}(M) \to \mathbb{R}$$

at each $\mathbf{x} \in U_{\mathbf{x}_0}$. So each $F(\mathbf{x}; \cdot, \cdot, \ldots, \cdot)$ has a different domain, and F is really a family of multilinear functionals parameterized by $\mathbf{x} \in M$. The proper treatment is to define F as a *section* (of a tensor product of vector bundles) as opposed to a function (with values in a tensor product of vector spaces). In fact, tensor fields are more than pointwise multilinear functionals, they satisfy the multilinearity condition in Eq. (6) with coefficients α, β being real-valued functions on M (usually in $C^\infty(M)$ if M is a smooth manifold) instead of merely being constants in $\mathbb{R}$.

The above discussions use the coordinate-free language of modern treatments of tensors and tensor fields in mathematics. In applications such as those considered in this survey, computations require introducing coordinates by chosing a basis on $\mathbb{V}$. If we pick a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$, where $n = \dim(\mathbb{V})$, then a multilinear functional f may be represented as an $n \times n \times \cdots \times n$ (p times) array of elements of $\mathbb{R}$:

$$\mathscr{A} = (a_{i_1 i_2 \cdots i_p})_{i_1, \ldots, i_p = 1}^n \in \mathbb{R}^{n \times \cdots \times n}. \tag{8}$$

We shall use the term *hypermatrix* of order p, or simply p-hypermatrix, when referring to a p-dimensional matrix of the form in Eq. (8). The origin of this terminology would appear to be [37]. These objects are natural multilinear generalizations of matrices in the following way. Since we have fixed a basis, every vector in $\mathbb{V}$ has a coordinate representation and we may assume that $\mathbb{V} = \mathbb{R}^n$. A bilinear functional $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ can be encoded by a matrix $\mathbf{A} = [a_{ij}]_{i,j=1}^n \in \mathbb{R}^{n \times n}$, in which the entry a_{ij} records the value of $f(\mathbf{e}_i, \mathbf{e}_j) \in \mathbb{R}$ where $\mathbf{e}_i$ denotes the ith standard basis vector in $\mathbb{R}^n$. By linearity in each coordinate, specifying $\mathbf{A}$ determines the values of f on all of $\mathbb{R}^n \times \mathbb{R}^n$; in fact, we have $f(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbf{A} \mathbf{v}$ for any (column) vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Thus, matrices encode all bilinear functionals. If $\mathbf{A} = \mathbf{A}^T$ is symmetric, the corresponding bilinear map is invariant under exchanging of coordinates:

$$f(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbf{A} \mathbf{v} = (\mathbf{u}^T \mathbf{A} \mathbf{v})^T = \mathbf{v}^T \mathbf{A}^T \mathbf{u} = \mathbf{v}^T \mathbf{A} \mathbf{u} = f(\mathbf{v}, \mathbf{u}).$$

To avoid sub-subscripts, we will restrict our discussion to 4-tensors. A 4-tensor is a quadrilinear functional $f : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ which has a coordinate representation given by a 4-hypermatrix $\mathscr{A} = (a_{ijkl})_{i,j,k,l=1}^n \in \mathbb{R}^{n \times n \times n \times n}$ as in Eq. (8) with $p = 4$. The subscripts and superscripts in Eq. (8) will be dropped whenever the

range of i, j, k, l is obvious or unimportant. A 4-hypermatrix is said to be *symmetric* if the value of a_{ijkl} stays the same for all 24 permutations of the indices:

$$a_{ijkl} = a_{ijlk} = a_{jilk} = \cdots = a_{lkji}.$$

Symmetric 4-tensors correspond to coordinate representations of quadrilinear maps $f : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with

$$f(\mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w}) = f(\mathbf{t}, \mathbf{u}, \mathbf{w}, \mathbf{v}) = f(\mathbf{u}, \mathbf{t}, \mathbf{v}, \mathbf{w}) = \cdots = f(\mathbf{w}, \mathbf{v}, \mathbf{u}, \mathbf{t}).$$

The set of symmetric 4-hypermatrices is often denoted $\mathsf{S}^4(\mathbb{R}^n)$ and it forms a linear subspace of the vector space $\mathbb{R}^{n \times n \times n \times n}$. More generally $\mathsf{S}^p(\mathbb{V})$, the set of symmetric p-tensors over an arbitrary vector space $\mathbb{V}$, may be defined in a coordinate-free manner [26] and forms a subspace of $\mathbb{V}^{\otimes p}$.

What about tensor fields? Since any manifold M may be given local coordinates, we may view tensor fields as hypermatrix-valued functions $F : M \to \mathbb{R}^{n \times \cdots \times n}$, $\mathbf{x} \mapsto \mathscr{A}_{\mathbf{x}} = (a_{i_1 i_2 \cdots i_p}(\mathbf{x}))_{i_1,\ldots,i_p=1}^n$, that are locally defined (roughly speaking, they are defined for local coordinates chosen for each neighborhood $U_{\mathbf{x}} \subseteq M$). The coordinate-dependent view of tensor fields as (hyper)matrix-valued functions is the classical approach. The subject, studied in this light, is often called tensor calculus, tensor analysis, or Ricci calculus. Tullio Levi-Civita, Gregorio Ricci-Curbastro, and Jan Schouten are usually credited for its invention [105].

3.2 Tensor Algebra and Homogeneous Polynomials

As we saw in the last section, a 4-hypermatrix $\mathscr{A} \in \mathbb{R}^{n \times n \times n \times n}$, is a coordinate representation of a 4-tensor, i.e., a quadrilinear functional $f : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. The set of 4-hypermatrices is naturally equipped with algebraic operations inherited from the algebraic structure of the tensor product space $\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n$:

- *Addition and Scalar Multiplication:* for $(a_{ijkl}), (b_{ijkl}) \in \mathbb{R}^{n \times n \times n \times n}$ and $\lambda, \mu \in \mathbb{R}$,

$$\lambda (a_{ijkl}) + \mu (b_{ijkl}) = (\lambda a_{ijkl} + \mu b_{ijkl}) \in \mathbb{R}^{n \times n \times n \times n}, \tag{9}$$

- *Outer Product Decomposition:* every $\mathscr{A} = (a_{ijkl}) \in \mathbb{R}^{n \times n \times n \times n}$ may be decomposed as

$$\mathscr{A} = \sum_{q=1}^r \lambda_q \mathbf{w}_q \otimes \mathbf{x}_q \otimes \mathbf{y}_q \otimes \mathbf{z}_q, \qquad a_{ijkl} = \sum_{q=1}^r \lambda_q w_{iq} x_{jq} y_{kq} z_{lq}, \tag{10}$$

with $\lambda_q \in \mathbb{R}$, $\mathbf{w}_q, \mathbf{x}_q, \mathbf{y}_q, \mathbf{z}_q \in \mathbb{R}^n$ for $q = 1, \ldots, r$. The symbol $\otimes$ here denotes the *Segre outer product*: for vectors $\mathbf{w} = [w_1, \ldots, w_n]^T, \ldots, \mathbf{z} = [z_1, \ldots, z_n]^T$,

$$\mathbf{w} \otimes \mathbf{x} \otimes \mathbf{y} \otimes \mathbf{z} := (w_i x_j y_k z_l)_{i,j,k,l=1}^n \in \mathbb{R}^{n \times n \times n \times n},$$

with obvious generalization to an arbitrary number of vectors. The ℓ-fold outer product of $\mathbf{x}$ with itself is written $\mathbf{x}^\ell$.

- *Multilinear Matrix Multiplication:* every $\mathscr{A} = (a_{ijkl}) \in \mathbb{R}^{n \times n \times n \times n}$ may be multiplied on its '4 sides' by matrices $\mathbf{W} = [w_{i\alpha}], \mathbf{X} = [x_{j\beta}], \mathbf{Y} = [y_{k\gamma}], \mathbf{Z} = [z_{l\delta}] \in \mathbb{R}^{n \times r}$ as follows

$$\mathscr{A} \cdot (\mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}) = (c_{\alpha\beta\gamma\delta})_{\alpha,\beta,\gamma,\delta=1}^{n} \in \mathbb{R}^{n \times n \times n \times n}, \qquad (11)$$

$$c_{\alpha\beta\gamma\delta} = \sum_{i,j,k,l=1}^{n} a_{ijkl} w_{i\alpha} x_{j\beta} y_{k\gamma} z_{l\delta}.$$

A different choice of bases $\mathbf{b}'_1, \ldots, \mathbf{b}'_n$ on $\mathbb{V}$ would lead to a different hypermatrix representation $\mathscr{B} \in \mathbb{R}^{n \times n \times n \times n}$ of elements in $\mathbb{V} \otimes \mathbb{V} \otimes \mathbb{V} \otimes \mathbb{V}$ — where the two hypermatrix representations $\mathscr{A}$ and $\mathscr{B}$ would be related precisely by a multilinear matrix multiplication of the form

$$\mathscr{A} \cdot (\mathbf{X}, \mathbf{X}, \mathbf{X}, \mathbf{X}) = \mathscr{B}$$

where $\mathbf{X}$ is the change-of-basis matrix, i.e. an invertible matrix with $\mathbf{X}\mathbf{b}_q = \mathbf{b}'_q$ for $q = 1, \ldots, n$. Therefore, a tensor and a hypermatrix are different in the same way a linear operator and a matrix are different. Note that in the context of matrices,

$$\mathbf{x} \otimes \mathbf{y} = \mathbf{x}\mathbf{y}^T \quad \text{and} \quad \mathbf{A} \cdot (\mathbf{X}, \mathbf{Y}) = \mathbf{Y}^T \mathbf{A} \mathbf{X}.$$

When $r = 1$ in Eq. (11), i.e., the matrices $\mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ are vectors $\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}$, we omit the $\cdot$ and write

$$\mathscr{A}(\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i,j,k,l=1}^{n} a_{ijkl} w_i x_j y_k z_l \qquad (12)$$

for the associated quadrilinear functional. Another special case occurs when one or more of the matrices $\mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ in Eq. (11) is the identity $\mathbf{I} = \mathbf{I}_{n \times n}$. For example,

$$\mathscr{A}(\mathbf{I}, \mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{j,k,l=1}^{n} a_{ijkl} x_j y_k z_l \in \mathbb{R}^n. \qquad (13)$$

In particular, the (partial) gradient of the quadrilinear functional $\mathscr{A}(\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z})$ may be expressed as

$$\nabla_{\mathbf{w}} \mathscr{A}(\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathscr{A}(\mathbf{I}, \mathbf{x}, \mathbf{y}, \mathbf{z}), \quad \nabla_{\mathbf{x}} \mathscr{A}(\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathscr{A}(\mathbf{w}, \mathbf{I}, \mathbf{y}, \mathbf{z}), \quad \text{etc.}$$

For a symmetric 4-tensor $\mathscr{S}$, we write $\mathscr{S} \cdot \mathbf{x}$ as a shorthand for $\mathscr{S}(\mathbf{x}, \mathbf{I}, \mathbf{I}, \mathbf{I})$; the result is a 3-tensor. Repeating this operation ℓ times is written $\mathscr{S} \cdot^\ell \mathbf{x}$. With this notation, the homogeneous quartic polynomial $\mathscr{S}(\mathbf{x})$ that is uniquely associated with $\mathscr{S}$ can be written as

$$\mathscr{S}(\mathbf{x}) := \mathscr{S}(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) = \mathscr{S} \cdot^4 \mathbf{x} = \sum_{d_1 + \cdots + d_n = 4} \mu_{d_1 \cdots d_n} \sigma_{d_1 \cdots d_n} x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}. \quad (14)$$

Similarly, the gradient of $\mathscr{S}(\mathbf{x})$ can be conveniently expressed as $\nabla \mathscr{S}(\mathbf{x}) = 4 \mathscr{S} \cdot^3 \mathbf{x}$. The right-hand side of Eq. (14) is the more typical way of writing a homogeneous polynomial in terms of monomials, unique coefficients $\sigma_{d_1, \ldots, d_n}$, and

multiplicities $\mu_{d_1,\ldots,d_n} := \binom{n}{d_1,\ldots,d_n}$. This is the higher-order equivalent of writing, for $\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$,

$$\mathbf{A}(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = ax_1^2 + bx_1x_2 + bx_2x_1 + cx_2^2 = ax_1^2 + 2bx_1x_2 + cx_2^2.$$

The *Frobenius norm* or *Hilbert-Schmidt norm* of a tensor $\mathscr{A}$ is defined by

$$\|\mathscr{A}\|_F^2 = \sum_{i,j,k,l=1}^n a_{ijkl}^2. \tag{15}$$

This is by far the most popular choice of norms used for a tensor since it is readily computable and also because it is induced by an inner product

$$\langle \mathscr{A}, \mathscr{B} \rangle = \sum_{i,j,k,l=1}^n a_{ijkl} b_{ijkl} \tag{16}$$

that generalizes the trace inner product. For symmetric p-tensors $\mathscr{S}, \mathscr{T}$ expressed in monomial form as in Eq. (14), this inner product may be written in the form

$$\langle \mathscr{S}, \mathscr{T} \rangle := \sum_{d_1+\cdots+d_n=p} \mu_{d_1\cdots d_n} \sigma_{d_1\cdots d_n} \tau_{d_1\cdots d_n}$$

and is often called the *apolar inner product* in invariant theory. For any $\mathbf{v} \in \mathbb{R}^n$, the apolar inner product of a symmetric tensor and the rank-1 symmetric tensor $\mathbf{v}^{\otimes p} := \mathbf{v} \otimes \cdots \otimes \mathbf{v}$ (p times),

$$\langle \mathscr{S}, \mathbf{v}^{\otimes p} \rangle = \mathscr{S}(\mathbf{v}),$$

which makes the set of symmetric p-tensors into a reproducing kernel Hilbert space.

3.3 Homogeneous Polynomials and Spherical Harmonics

By restricting Eq. (14) to the 3D unit sphere S^2, $\mathbf{x} = [\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta]^T$, every symmetric tensor $\mathscr{S}$ defines a real-valued homogeneous polynomial function on S^2. Spherical harmonics (SH) are an alternate basis for describing functions on the sphere. The SHs form a complex complete orthonormal basis for square integrable functions on the unit sphere. Spherical functions can, therefore, be naturally expanded in the infinite SH basis or approximated to any accuracy by a truncated series. Again the diffusion signal being real and symmetric, a modified real and symmetric SH basis is chosen in dMRI. Therefore, S can be written as

$$S(\theta, \phi) = \sum_{j=1}^{M'} c_j Y_j(\theta, \phi), \tag{17}$$

where $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$ and c_j are the coefficients describing S in the modified SH basis [30]

$$Y_j(\theta,\phi) = \begin{cases} \sqrt{2}\,\mathrm{Re}(Y_l^{|m|}(\theta,\phi)) & \text{if } m < 0, \\ Y_l^m(\theta,\phi) & \text{if } m = 0, \\ (-1)^{m+1}\sqrt{2}\,\mathrm{Im}(Y_l^m(\theta,\phi)) & \text{if } m > 0, \end{cases} \tag{18}$$

with $Y_l^m(\theta,\phi)$ the rank l and degree m regular complex spherical harmonic:

$$Y_l^m(\theta,\phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos\theta)e^{im\phi}, \quad m \le |l|. \tag{19}$$

In [86, 29] it was shown that the tensor basis and the SH basis are bijective via a linear transformation when the rank l of the truncated SH basis equals the order k of the symmetric tensor. This can be understood from the spherical harmonic transform of the polynomial representation of S:

$$c_j = \sum_{i=1}^{M'=M} \mu_i \sigma_i \int_{S_2} x_1^{\alpha_i} x_2^{\beta_i} x_3^{l-\alpha_i-\beta_i} Y_j(\theta,\phi)d\Omega. \tag{20}$$

where the new indexing of μ and σ assumes an arbitrary ordering of the $\mu_{d_1\cdots d_n}$ and $\sigma_{d_1\cdots d_n}$ from Eq. (14). Since the integral does not depend on the tensor coefficients σ_j, Eq. (20) can be seen as a dot product between the vector of unique tensor coefficients and the vector of spherical harmonic transforms of the M monomials $x_1^{\alpha_i} x_2^{\beta_i} x_3^{l-\alpha_i-\beta_i}$. In other words, computing the M SH coefficients can be written as a matrix vector multiplication

$$\mathbf{c} = \mathbf{Ms}, \tag{21}$$

where $\mathbf{c} = [c_1, c_2, \ldots, c_M]^T$, $\mathbf{s} = [\sigma_1, \sigma_2, \ldots, \sigma_M]^T$, and:

$$\mathbf{M} = \begin{bmatrix} \mu_1 \int_{S_2} x_1^{\alpha_1} x_2^{\beta_1} x_3^{l-\alpha_1-\beta_1} Y_1 d\Omega & \cdots & \mu_M \int_{S_2} x_1^{\alpha_M} x_2^{\beta_M} x_3^{l-\alpha_M-\beta_M} Y_1 d\Omega \\ \vdots & \ddots & \vdots \\ \mu_1 \int_{S_2} x_1^{\alpha_1} x_2^{\beta_1} x_3^{l-\alpha_1-\beta_1} Y_M d\Omega & \cdots & \mu_M \int_{S_2} x_1^{\alpha_M} x_2^{\beta_M} x_3^{l-\alpha_M-\beta_M} Y_M d\Omega \end{bmatrix}. \tag{22}$$

3.4 Tensor Decompositions and Approximations

A tensor that can be expressed as an outer product of vectors is called *decomposable* and rank-1 if it is also nonzero. More generally, the *rank* of a tensor $\mathscr{A} = (a_{ijkl})_{i,j,k,l=1}^n \in \mathbb{R}^{n \times n \times n \times n}$, denoted rank$(\mathscr{A})$, is defined as the minimum r for which $\mathscr{A}$ may be expressed as a sum of r rank-1 tensors [52, 53],

$$\mathrm{rank}(\mathscr{A}) := \min\left\{ r \,\Big|\, \mathscr{A} = \sum_{q=1}^r \lambda_q \mathbf{w}_q \otimes \mathbf{x}_q \otimes \mathbf{y}_q \otimes \mathbf{z}_q \right\} \tag{23}$$

where the minimum is taken over all decomposition with $\lambda_p \in \mathbb{R}$, $\mathbf{w}_p, \mathbf{x}_p, \mathbf{y}_p, \mathbf{z}_p \in \mathbb{R}^n$, $p = 1, \ldots, r$. If $\mathscr{S}$ is a symmetric tensor, then its *symmetric rank* [26] is

$$\text{srank}(\mathscr{S}) := \min\left\{ r \,\middle|\, \mathscr{S} = \sum_{q=1}^{r} \lambda_q \mathbf{x}_q \otimes \mathbf{x}_q \otimes \mathbf{x}_q \otimes \mathbf{x}_q \right\}. \tag{24}$$

We remark that it is not known whether the rank of a symmetric tensor is equal to its symmetric rank. The definition of rank in Eq. (23) agrees with matrix rank when applied to an order-2 tensor. In certain other literature, for example [90], the term 'rank' is used synonymously with what we called 'order' in the first paragraph of this section. For tensors of order greater than 2, rank becomes a more intricate notion than matrix rank with properties that may seem surprising at first encounter. We refer the readers to [28] (rank) and [26] (symmetric rank) for further information.

Best rank-r approximations

$$\text{argmin}_{\lambda \in \mathbb{R}^r, \, \mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathbb{R}^{n \times r}} \left\| \mathscr{A} - \sum_{q=1}^{r} \lambda_q \mathbf{w}_q \otimes \mathbf{x}_q \otimes \mathbf{y}_q \otimes \mathbf{z}_q \right\| \tag{25}$$

and the corresponding best symmetric rank-r approximation problem (i.e. when $\mathbf{W} = \mathbf{X} = \mathbf{Y} = \mathbf{Z}$) are used in practice (Section 5.4), but have no solution in general when $r > 1$. The easiest way to explain this is that the *infimum* of the objective function, taken over all $\lambda = (\lambda_1, \ldots, \lambda_r) \in \mathbb{R}^r$ and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$, $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_r]$, $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_r]$, $\mathbf{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_r] \in \mathbb{R}^{n \times r}$ need not be attained. This happens regardless of symmetry, the choice of norms in Eq. (25) and for any order $p \geq 3$. In the unsymmetric case, it is known that the set of tensors of rank $s > r$ that do not have a best rank-r approximation could form a positive volume set. A particularly eggregious case is $\mathbb{R}^{2 \times 2 \times 2}$, where no rank-3 tensor has a best rank-2 approximation. Fortunately, there are special cases where the problem can be alleviated, notably: (i) when all coordinates of $\mathscr{A}$ are nonnegative and $\lambda, \mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z} \geq 0$ [74]; (ii) when $\mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ satisfy a 'coherence' condition [75]; (iii) when p is even and $\lambda \geq 0$ [76]. Unlike cases (i) and (ii), case (iii) only applies to symmetric approximations.

3.5 Eigenvectors and Eigenvalues

The basic notions for eigenvalues of tensors were introduced independently by Lim [72] and Qi [93]. The usual eigenvalues and eigenvectors of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ are the stationary values and stationary points of its Rayleigh quotient, and this point of view generalizes naturally to tensors of higher order. This gives, for example, an *eigenvector* of a tensor $\mathscr{A} = (a_{ijkl})_{i,j,k,l=1}^{n} \in \mathbb{R}^{n \times n \times n \times n}$ as a nonzero column vector $\mathbf{x} = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$ satisfying

$$\sum_{i,j,k=1}^{n} a_{ijkl} x_i x_j x_k = \lambda x_l, \quad l = 1, \ldots, n, \tag{26}$$

for some $\lambda \in \mathbb{R}$, which is called an *eigenvalue* of $\mathscr{A}$. Notice that if $(\lambda, \mathbf{x})$ is an eigenpair, then so is $(t\lambda, t\mathbf{x})$ for any $t \neq 0$; thus, eigenpairs are more naturally defined projectively. As in the matrix case, generic tensors over $\mathbb{R}$ or $\mathbb{C}$ have a finite number of eigenvalues and eigenvectors (up to this scaling equivalence), although their count is exponential in n. Still, it is possible for a tensor to have an infinite number of

eigenvalues, but in that case they comprise a cofinite set of complex numbers. For an even-ordered symmetric tensor $\mathscr{S} \in \mathsf{S}^{2p}(\mathbb{R}^n)$, one has that $\mathscr{S}$ is nonnegative definite, i.e. $\mathscr{S}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, if and only if all the eigenvalues of $\mathscr{S}$ are nonnegative [93] — a generalization of a well-known fact for symmetric tensors.

It is worth noting that unlike in the matrix case, most tensor problems are NP-hard. This includes determining rank, best rank-1 approximation, spectral norm, eigenvalues, and eigenvectors [51]. However, the notion of NP-hardness is an asymptotic one that applies when $n \to \infty$. Therefore, these hardness results do not preclude the existence of efficient algorithms for a fixed n, and especially for small values such as $n = 3$, the case of greatest interest to diffusion MRI.

4 Fitting Higher-Order Tensor Models

4.1 Fitting Models of Apparent Diffusivity

One of the earliest models that attempted to overcome the limitations of second-order diffusion tensors used HOTs to account for diffusion with generalized angular profiles while preserving its radial monoexponential behaviour [86]. Even order Cartesian tensors were used to measure the apparent diffusivities (ADC) from the generalized Stejskal-Tanner equation as described in Eq. (2).

The simplest method [86] for estimating such tensors, $\mathscr{D}$, from the diffusion signal is to linearize the Stejskal-Tanner equation by taking the logarithm of Eq. (2). This leads to system of linear equations: $\mathbf{A}\mathbf{x} = \mathbf{y}$, where the rows of the design matrix $\mathbf{A}$ contain the monomials of the homogeneous form $D(\mathbf{u}) = \mathscr{D}^{(k)} \cdot^k \mathbf{u}$, the vector $\mathbf{y}$ contains the log-normalized diffusion signal scaled by the acquisition parameter b, and the vector $\mathbf{x}$ contains the unknown coefficients of the tensor $\mathscr{D}$. This system is overdetermined when the number of data acquisitions is greater than the number of unknown tensor coefficients and can be solved uniquely in the least squares sense by taking the Moore-Penrose pseudo-inverse of $\mathbf{A}$.

Since diffusivity is a non-negative physical quantity, the homogeneous form $D(\mathbf{u})$ cannot be negative for any $\mathbf{u} \in S^2$. This leads to a positivity constraint that needs to be respected while estimating $\mathscr{D}$. The least squares approach often violates this constraint for $\mathscr{D}$ with high orders and when the acquisitions are noisy.

Descoteaux et al. [29] proposed a linear approach with angular regularization to account for noisy acquisitions. Leveraging the bijection between HOTs and SHs, see Eq. (21), they estimated the coefficients of $\mathscr{D}$ by first estimating the coefficients in an SH basis of rank equal to the order of the tensor while applying Laplace-Beltrami smoothing on the sphere and then converting back to the tensor basis. This again leads to a linear system that is overdetermined when the number of acquisitions is larger than the number of tensor coefficients,

$$\mathbf{x} = \mathbf{M}^{-1}(\mathbf{B}^T\mathbf{B} + \lambda\mathbf{L})^{-1}\mathbf{B}^T\mathbf{y}, \tag{27}$$

where **x** contains the unique tensor coefficients, **y** contains the log-normalized signal, **B** is the design matrix in the SH basis and **M** represents the linear transformation matrix between the HOT basis and the SH basis. The matrix **L** is a diagonal matrix with entries $l_{ii} = \ell_i^2(\ell_i + 1)^2$, which represents the Laplace-Beltrami regularization of the SH Y_ℓ^m, and λ is the regularization weight. This becomes the least squares solution when $\lambda = 0$, but with nonzero λ, **L** tends to smooth higher order terms more, therefore, dampening the effects of noise in higher orders.

Florack et al. used the same Laplace-Beltrami regularization on the sphere, but for tensors instead of SHs [34]. This was based on an infinite inhomogeneous tensor basis representation, much like the SHs, with the diffusion function modified to $\tilde{D}(\mathbf{u}) = \sum_{k=0}^{\infty} \mathscr{D}^{(k)} \cdot^k \mathbf{u}$. It was shown that on the sphere, this representation was redundant, and when truncated to a finite order, it represented the same diffusion function as in Eq. (2). The relation between the homogeneous and inhomogeneous tensor representation has been addressed rigorously in [8]. The estimation process was specifically crafted such that higher order tensors only captured the residual information not available in lower order tensors. This resulted in a "canonical" tensorial representation where the span of a tensor of fixed order k formed a degenerate eigenspace for the Laplace-Beltrami operator with eigenvalue $-k(k+1)$, exactly like the SHs.

The problem of estimating $\mathscr{D}$ with the positivity constraint was solved for order 4 tensors, in two different ways. The homogeneous forms of symmetric order 4 tensors of dimension 3 are known as ternary quartics. Barmpoutis et al. [10, 9] and Ghosh et al. [43] use Hilbert's theorem on positive semi-definite (psd) ternary quartics:

Theorem 1. *If $P(x, y, z)$ is homogeneous, of degree 4, with real coefficients and $P(x, y, z) \geq 0$ at every $(x, y, z) \in \mathbb{R}^3$, then there are quadratic homogeneous polynomials f, g, h with real coefficients, such that $P = f^2 + g^2 + h^2$.*

Therefore, estimating $P(x, y, z)$ (or $\mathscr{D}^{(4)}$) by estimating f, g, h ensures $\mathscr{D}^{(4)}$ to be psd. However, these quadratic polynomials can only be uniquely determined up to a 3D rotation and up to a sign. In other words, if the 6 coefficients of f, g, h each are written as column vectors $\mathbf{w}_f, \mathbf{w}_g, \mathbf{w}_h$, respectively, and a 6×3 matrix $\mathbf{W} = [\mathbf{w}_f, \mathbf{w}_g, \mathbf{w}_h]$ is constructed, then $P(x, y, z) = \mathbf{v}^T \mathbf{W} \mathbf{W}^T \mathbf{v}$, where $\mathbf{v}^T = [x^2, y^2, z^2, xy, xz, yz]$. Thus $\mathbf{W}$, $-\mathbf{W}$ and $\mathbf{W}\mathbf{R}$ for any 3×3 orthogonal matrix $\mathbf{R}$ result in the same P.

Initially, Barmpoutis et al. fixed $\mathbf{R}$ by choosing the rotation that renders $\mathbf{A}$ – the top 3×3 block of $\mathbf{W}$ – to a lower triangular matrix [10]. This was achieved by considering the QR-decomposition of $\mathbf{A}$, but in practice $\mathbf{A}$ was taken to be lower triangular. This resulted in a reduction of unknown coefficients from $18 = 3 \times 6$ to 15, which is exactly the number of unique coefficients of $\mathscr{D}^{(4)}$. In a later work [9], an Iwasawa decomposition of $\mathbf{W}\mathbf{W}^T$ was taken, which implied the Cholesky decomposition of $\mathbf{A}$. This again resulted in $\mathbf{A}$ being rendered lower triangular – defining uniqueness over 3D rotations and again reducing the number of unknowns to 15. Furthermore, the Cholesky decomposition constrained the diagonal entries of $\mathbf{A}$ to be positive – defining uniqueness over the sign.

Ghosh et al. [43] estimated all 18 unknowns of $\mathbf{W}$ and reconstructed the 15 coefficients of $\mathscr{D}^{(4)}$ from the Gram matrix $\mathbf{W}\mathbf{W}^T$. Although $\mathbf{W}$ cannot be estimated

uniquely, the Gram matrix representing the homogeneous form P is unique and the mapping from the coefficients of the Gram matrix to the coefficients of $\mathscr{D}^{(4)}$ is unique. Therefore, the estimation of the tensor coefficients is unambiguous. While Barmpoutis et al. [10, 9] used a Levenberg-Marquardt optimization scheme, Ghosh et al. [43] prefer the Broyden-Fletcher-Goldfarb-Shanno (BFGS) scheme.

Barmpoutis et al. [10, 9] further introduced an *L2* distance measure between the homogeneous forms corresponding to the tensors evaluated on the unit sphere

$$\text{dist}(\mathscr{D}_1^{(4)}, \mathscr{D}_2^{(4)})^2 = \frac{1}{4\pi} \int_{S^2} [D_1(\mathbf{u}) - D_2(\mathbf{u})]^2 d\mathbf{u}, \tag{28}$$

which was computed analytically in terms of the difference of the coefficients of $\mathscr{D}_1^{(4)}$ and $\mathscr{D}_2^{(4)}$, and which was used for spatial regularization of the tensor field to account for noise.

A second way of estimating $\mathscr{D}^{(4)}$ with the positivity constraint was proposed by Ghosh et al. [44]. In this approach, the 6×6 isometrically equivalent matrix representation [20] $\mathbf{D}$ of $\mathscr{D}^{(4)}$ was used. Since $\mathbf{D}$ is symmetric and its positive definiteness ensures $\mathscr{D}^{(4)}$ to be positive, the affine invariant Riemannian metric for the space of symmetric positive definite matrices [71] was used to estimate $\mathbf{D}$ via a Riemannian gradient descent. However, the symmetry of the tensor $\mathscr{D}^{(4)}$ cannot be entirely captured by $\mathbf{D}$, which has 21 unique coefficients. Therefore, a final symmetrizing step was used to recover a positive and symmetric tensor $\mathscr{D}^{(4)}$.

The problem of estimating an arbitrary even order HOT, $\mathscr{D}^{(2k)}$, with the positivity constraint was also solved in two different ways. Barmpoutis et al. [13] used a result that states that for any even degree, $2k$, a (homogeneous) polynomial positive on the unit sphere can be written as a sum of squares of polynomials, p, of degree k on the unit sphere, $D(\mathbf{u}) = \mathscr{D}^{(2k)} \cdot^{(2k)} \mathbf{u} = \sum_{j=1}^{R} \lambda_j p^{(k)}(\mathbf{u}, \mathbf{c}_j)^2$, where λ_j are all positive and $\mathbf{c}_j$ are the coefficient vectors of the polynomials p_j with $||\mathbf{c}_j|| = 1$. However, since R, the number of polynomials in the sum, cannot be determined, they reformulated the problem as a spherical convolution problem $D(\mathbf{u}) = \int_{S^{\#\mathbf{c}-1}} \lambda(\mathbf{c}) p^{(k)}(\mathbf{u}; \mathbf{c})^2 d\mathbf{c}$, where the unit sphere $S^{\#\mathbf{c}-1}$ is embedded in $\mathbb{R}^{\#\mathbf{c}}$, with #**c** being the number of elements in **c**. The convolution was solved numerically by discretizing $S^{\#\mathbf{c}-1}$ finely and $\mathscr{D}^{(2k)}$ was estimated by solving the least squares problem for the unknowns λ_j

$$E = \sum_{i=1}^{N} \left(S_i/S_0 - e^{-b\sum_{j=1}^{r} \lambda_j p(\mathbf{g}_i; \mathbf{c}_j)^2} \right)^2 \tag{29}$$

using a non-negative least squares (NNLS) to ensure that all $\lambda_j \geq 0$. Eq. (29) essentially overestimates R by r by discretizing the convolution, while the NNLS tends to compute a sparse solution ensuring that Eq. (29) does not overfit the signal.

A second method for estimating even order psd HOTs based on convex optimization was proposed by Qi et al. [96]. It was shown that the set of order $2k$ psd HOTs, $\mathscr{D}$, form a closed convex cone $\mathscr{C}$ in $\mathbb{R}^n$, where $\mathscr{D}$ has n unique coefficients and can be represented by $\mathbf{x} \in \mathbb{R}^n$. Furthermore, the psd constrained least squares estimation was shown to be convex and quadratic with a unique minimizer $\mathbf{x}^* \in \mathscr{C}$ such that if the unconstrained solution $\bar{\mathbf{x}} \in \mathbb{R}^n \setminus \mathscr{C}$ then $\mathbf{x}^* \in \partial \mathscr{C}$, the boundary of $\mathscr{C}$.

The explicit psd constraint on $\mathscr{D}$ was formulated as $\lambda_{\min}(\mathscr{D}) \geq 0$ where $\lambda_{\min}(\mathscr{D})$, the minimum Z-eigenvalue of $\mathscr{D}$, was shown to be computationally tractable. The psd HOT $\mathscr{D}_{\mathbf{x}^*}$ (corresponding to $\mathbf{x}^*$) was estimated by first checking the psd-ness of the unconstrained HOT $\mathscr{D}_{\bar{\mathbf{x}}}$. If $\lambda_{\min}(\mathscr{D}_{\bar{\mathbf{x}}}) \geq 0$, then by uniqueness $\mathscr{D}_{\mathbf{x}^*} = \mathscr{D}_{\bar{\mathbf{x}}}$. However, if $\bar{\mathbf{x}} \notin \mathscr{C}$, then $\mathscr{D}_{\mathbf{x}^*}$ was estimated by solving the non-differentiable, *non-convex* optimization problem $L(\mathbf{x}) = \min\{|\mathbf{Ax} - \mathbf{y}|^2 : \lambda_{\min}(\mathscr{D}_{\mathbf{x}}) = 0\}$, with only an equality constraint, by a subgradient descent approach. In theory, $\mathscr{D}_{\mathbf{x}^*}$ could also be estimated by solving the psd constrained non-differentiable *convex* least squares problem.

Alternatively, Barmpoutis et al. [11] used even ordered HOTs to model the logarithm of the diffusivities. This preserved the monoexponential radial diffusion but considered the exponential of the tensor for the angular diffusion $D(\mathbf{u}) = \exp(\mathscr{D}^{(k)} \cdot^k \mathbf{u})$ (in Eq. (2)). This automatically ensured positive diffusion without having to impose any constraints. The approach was inspired by the Log-Euclidean metric for DTI [3].

4.2 Fitting Models of Apparent Diffusional Kurtosis

Fitting the coefficients of the diffusion tensor $\mathbf{D}$ and kurtosis tensor $\mathscr{W}$ in Eq. (5) is simplified by initially considering each gradient direction separately, and finding parameters of the corresponding one-dimensional diffusion process,

$$S(b) = S_0 \mathrm{e}^{-bd + \frac{1}{6}(bd)^2 K},$$ (30)

where d and K are apparent diffusion and kurtosis coefficients, respectively. Estimating these two variables requires measurements $S(b)$ with at least two non-zero b-values, in addition to the baseline S_0 measurement. After taking the logarithm on both sides of Eq. (30), this leads to a system of equations that is quadratic in d, and can thus no longer be solved with a linear least squares estimator. Instead, gradient-based iterative Levenberg-Marquardt optimization has been employed [61].

Assuming a Gaussian noise model results in a positive bias in the estimated kurtosis values, which can be removed by finding the maximum likelihood fit under a Rician noise model [112] or, more easily, by accounting for the noise-induced bias in the measurements themselves [61, 82]. This is done by adding an estimate η of the background noise to the signal model in Eq. (30),

$$S(b) = \sqrt{\eta^2 + \left(S_0 \mathrm{e}^{-bd + \frac{1}{6}(bd)^2 K}\right)^2}.$$ (31)

After finding parameters d_i and K_i for each individual gradient direction i, a second-order diffusion tensor $\mathbf{D}$ can be fit linearly to the d_i. Given this estimate of $\mathbf{D}$, the fourth-order kurtosis tensor $\mathscr{W}$ can then be fit linearly using Eq. (4) [82, 69].

Kurtosis is a dimensionless quantity and can, in theory, take on any value $K \geq -2$. However, the kurtosis of a system that contains noninteracting Gaussian compart-

ments with different diffusivities is always non-negative, and empirical results suggest non-negative kurtosis in human brain tissue [61]. Similarly, an upper bound on kurtosis, $K_i \leq 3/(b_{max}d_i)$, where b_{max} is the largest b-value used in the measurements, is implied by the empirical observation that in practice, the signal $S(b)$ is a monotonically decreasing function of b. These two constraints have been enforced as part of the fitting, using quadratic programming or heuristic thresholding [106]. Other authors have chosen to merely enforce the lower bound $K \geq -3/7$, which correspond to the kurtosis of water confined to equally-sized spherical pores, by a sum-of-squares parametrization of the homogeneous polynomial represented by $\mathscr{W}$ [16]. Additional regularization has been employed to penalize extrema in the homogeneous form that fall outside the range of the measured kurtosis values [65].

4.3 Fitting Deconvolution-Based Models

Spherical deconvolution models the diffusion-weighted signal $S(\mathbf{u})$ in different gradient directions $\mathbf{u}$ as the convolution of a fiber orientation density function (fODF) F with a response function R. It describes the signal attenuation caused by a single nerve fiber bundle, and it is assumed to be cylindrically symmetric:

$$S(\mathbf{u}) = \iint_{\|\mathbf{v}\|=1} F(\mathbf{v})\, R(\mathbf{v} \cdot \mathbf{u})\, d\mathbf{v} \tag{32}$$

Based on Eq. (32), deconvolution can be used to estimate the fiber ODF F from the measurements S. Deconvolution is done most easily in the spherical harmonics basis, where it amounts to simple scalar division. However, constructing a spherical harmonics representation of the deconvolution kernel R requires two choices: Beside estimating the response of a single fiber compartment from the data [108] or deriving it from an analytical fiber model [31, 102], it involves deciding how the single fiber compartment should be represented *after* deconvolution [107].

Even though the delta distribution may seem like an obvious choice, it requires an infinite number of coefficients in the spherical harmonics basis. Therefore, Tournier et al. [107] approximate the delta peak, resulting in non-trivial interactions between peaks of non-orthogonal fiber compartments and leading to systematic errors when taking ODF maxima as estimates of fiber directions, even when no measurement noise is present. Schultz and Seidel [101] have removed this problem by instead modeling single fiber peaks as rank-1 tensors, and performing a low-rank approximation of the resulting order-p fODF tensor $\mathscr{F}$,

$$\underset{\lambda_i, \mathbf{v}_i}{\operatorname{argmin}} \left\| \mathscr{F} - \sum_{i=1}^{r} \lambda_i \mathbf{v}_i^{\otimes p} \right\|_F, \tag{33}$$

where $\mathbf{v}_i$ describe the per-compartment principal directions, and λ_i are proportional to their volume fractions. The approximation rank r corresponds to the number of

discrete fiber compartments; one way to estimate it is by learning from simulated training data via support vector regression [98].

This tensor-based variant of spherical deconvolution uses the linear bijection between spherical harmonics and polynomial bases (cf. Section 3.3) twice: First, to map a rank-1 tensor of the same order as the desired fODF tensor $\mathscr{F}$ to the spherical harmonics basis, which is required to find the correct kernel **R** for use with that tensor order. Second, to transform the deconvolution result F, obtained in the spherical harmonics basis, back into its tensor representation $\mathscr{F}$.

Since compartments cannot have negative weights, valid fODF tensors should permit a positive decomposition into rank-1 terms. For tensor order $k > 2$, this is a stronger requirement than non-negativity of the homogeneous form, which is a more natural constraint for models of apparent diffusivity (Section 4.1). It can be enforced by computing an approximation with the generic number of rank-1 terms and non-negative weights [99].

Similar to a previous approach of Barmpoutis et al. [13], Weldeselassie et al. [114] enforce non-negativity of F by parametrizing the homogeneous polynomial $\mathscr{F}$ (with even order k) as a sum of squares of polynomials of order $k/2$. Rather than performing the deconvolution in spherical harmonics, they discretize the fODF, so that it can be found as the non-negative least squares solution of a linear system.

4.4 Fitting Other Types of Models

When fitting the higher-order diffusion model described by Eq. (3) [78, 77], we only consider tensors of even order, as was argued in Section 2. By taking the logarithm and truncating after order $2n$, the equation can be rewritten in the form

$$\mathrm{Re}[\log(S(\mathscr{B})/S_0)] = \sum_{k=1}^{n}(-1)^k\langle\mathscr{B}^{(2k)},\mathscr{D}^{(2k)}\rangle,$$

where Re denotes the real part of the logarithmic signal and the inner products between tensors $\mathscr{B}^{(2k)}$ and $\mathscr{D}^{(2k)}$ is defined in Eq. (16). Tensors $\mathscr{D}^{(2k)}$ can be estimated by considering measurements with different gradient strengths and directions, which lead to different $\mathscr{B}_i^{(2k)}$, and truncating the tensor series at the desired order. If we have m measurements, we obtain m equations of the above form, linear in the coefficients of $\mathscr{D}$. These can be combined in a matrix equation

$$\mathbf{y}(\log(|S_i(\mathscr{B}_i)|/|S_0|)) = \mathbf{B}(\mathscr{B}^{(2k),i})\,\mathbf{x}(\mathscr{D}^{(2k)}), \tag{34}$$

where $i = 1,\ldots,m$. In practice, the modulus $|\cdot|$ rather than the real part of the complex signal is used, since phase is unreliable. The vector **x**, which contains the coefficients of $\mathscr{D}^{(2k)}$, can be estimated by solving Eq. (34) in the least squares sense.

Higher-order tensors representing q-ball ODFs (see Section 2) can also be fitted to HARDI data. An analytical solution for the q-ball ODF is given by [2, 50, 30]

$$\psi_{\text{q-ball}}(\mathbf{u}) = \sum_{i=1}^{N} 2\pi P_{l_i}(0)\, c_i\, Y_i(\mathbf{u}) \tag{35}$$

where $\mathbf{u}$ is a unit norm vector, P_{l_i} is the Legendre polynomial of degree l_i, $\{Y_i\}_{i=1}^{N}$ is a modified SH basis as in Eq. (18), and c_i are the harmonic coefficients of the MR signal. A tensor representation of $\psi_{\text{q-ball}}$ can be obtained from the bijection between SHs and tensors. Alternatively, it can be reconstructed directly in a tensor basis [34]

$$\psi_{\text{q-ball}}(\mathbf{u}) = \sum_{k=0}^{n} 2\pi P_{l_k}(0)\mathscr{S}_k \cdot^{k} \mathbf{u} \tag{36}$$

where n is the maximum order of a series of tensors $\mathscr{S}_k$ fitted to the diffusion signal such that higher orders only encode the fitting residuals from lower orders.

5 Processing Higher-Order Tensors in Diffusion MRI

5.1 Computing Rotationally Invariant Scalar Measures

It is desirable to extract meaningful scalars from the estimated higher-order tensors. In particular, rotationally invariant quantities are preferable. These are independent of the coordinate system and thus intrinsic features of the tensor.

5.1.1 Higher-Order Diffusion Tensors

Rotationally invariant measures of diffusivity and anisotropy based on higher-order diffusion tensors have been proposed in [89]. The mean diffusivity is defined as:

$$\langle D \rangle = \frac{1}{4\pi} \iint_{S^2} D(\mathbf{u})\, d\mathbf{u} \tag{37}$$

where $\mathbf{u}$ is a unit direction vector and $D(\mathbf{u})$ are the diffusivities as in Eq. (2). The generalized anisotropy (GA) and scaled entropy (SE) are given by

$$\text{GA} = 1 - \frac{1}{1 + (250\text{V})^{\varepsilon(V)}} \quad \text{and} \quad \text{SE} = 1 - \frac{1}{1 + (60(\ln 3 - \eta))^{\varepsilon(\ln 3 - \eta)}}, \tag{38}$$

where $\varepsilon(\gamma) = 1 + 1/(1 + 5000 \cdot \gamma)$ and V and η are the variance and entropy of the normalized diffusivities, $D(\mathbf{u})/3\langle D \rangle$. The definition of these measures does not rely on any specific tensor order. In addition, GA and SE are scaled between 0 and 1. Note that these measures can also be calculated from other functions defined on the unit sphere, such as orientation distribution functions.

GA and SE values for simulated data modelling two and three fibers show a clear difference between those implied by tensors of order two and higher-order (four, six and eight) tensors, the latter being significantly higher [89, 29, 84, 27]. GA and SE have also been reported to be slightly higher in the case of sixth order tensors than

for order four [27]. On the other hand, for data simulating one fiber, GA and SE are independent of the tensor order. This is also the case for the mean diffusivity [89].

GA and SE for real HARDI data of healthy subjects have been studied in [84, 27, 83]. It has been shown that fourth- and sixth-order tensors result in increased values for both measures, especially for SE, with respect to second-order tensors. This effect is observed in areas with intra-voxel orientational heterogeneity but also in some regions with coherent axonal orientation. On the other hand, GA and SE become more sensitive to noise for increasing tensor order [27].

The variance of fourth-order covariance tensors has also been investigated for DTI data of glioblastoma patients [32]. Results indicate a better variance contrast between tumor subregions than for FA.

5.1.2 Diffusional Kurtosis Tensors

A number of rotationally invariant scalar measures based on fourth-order kurtosis tensors have been proposed. Different definitions of mean kurtosis (also referred to as average AKC), kurtosis anisotropy, radial and axial kurtoses can be found in the literature. Some of them are related to certain eigenvalues of the kurtosis tensor, which we discuss later in this section. These measures are summarized in Tables 1 and 2. It is clear that they are rotationally invariant, since both the AKC and eigenvalues involved in their definition are rotationally invariant.

Table 1 Mean kurtosis and kurtosis anisotropy. β: D-eigenvalue of $\mathscr{W}$, v: number of D-eigenvalues, N: total number of diffusion measurement directions, K_{app}: AKC in a particular direction as in Eq. (5), $\mathbf{e_i}$ ($i = 1, 2, 3$): eigenvector of diffusion tensor $\mathscr{D}$, $K_{\mathrm{app}}(\mathbf{e_i}) = (\mathrm{MD}^2/\lambda_i^2) \cdot \mathscr{W}_{iiii}$, $\mathscr{W}$: kurtosis tensor in the basis $\{\mathbf{e_i}\}$, $\bar{K} = (1/3)(K_{\mathrm{app}}(\mathbf{e_1}) + K_{\mathrm{app}}(\mathbf{e_2}) + K_{\mathrm{app}}(\mathbf{e_3}))$.

Reference	MK	FA$_K$
[95]	$\frac{(\mathrm{MD})^2}{v} \sum_{i=1}^{v} \beta_i$	$\sqrt{\frac{v}{v-1}} \sqrt{\frac{\sum_{i=1}^{v} (\beta_i - \mathrm{MK}/(\mathrm{MD})^2)^2}{\sum_{i=1}^{v} \beta_i^2}}$
[56]	$\frac{1}{N} \sum_{i=1}^{N} K_{\mathrm{app}}(\mathbf{u}_i)$	$\sqrt{\frac{3}{2}} \sqrt{\frac{\sum_{i=1}^{3} (K_{\mathrm{app}}(\mathbf{e_i}) - \bar{K})^2}{\sum_{i=1}^{3} K_{\mathrm{app}}^2(\mathbf{e_i})}}$
[94, 106, 60] (MK) [91]	$\frac{1}{4\pi} \iint_{\mathrm{S}^2} K_{\mathrm{app}}(\mathbf{u})\, d\mathbf{u}$	$\sqrt{\frac{1}{4\pi} \iint_{\mathrm{S}^2} (K_{\mathrm{app}}(\mathbf{u}) - \mathrm{MK})^2\, d\mathbf{u}}$
[81]	idem	$\sqrt{\frac{\iint_{\mathrm{S}^2} (K_{\mathrm{app}}(\mathbf{u}) - \mathrm{MK})^2\, d\mathbf{u}}{\iint_{\mathrm{S}^2} K_{\mathrm{app}}^2(\mathbf{u})\, d\mathbf{u}}}$

Note that the first two definitions of kurtosis anisotropy in Table 1 are completely analogous to the DTI case but based on the kurtosis tensor D-eigenvalues and AKC

values along the diffusion tensor eigenvectors, respectively. As FA, FA_K takes on values $0 \leq \text{FA}_K \leq 1$, except for the definition in [91].

Table 2 Axial and radial kurtoses. $\mathbf{e}_\phi = (0, \cos\phi, \sin\phi)$ in the basis $\{\mathbf{e_i}\}$.

Reference	$K_\parallel$	$K_\perp$
[56]	$K_{\text{app}}(\mathbf{e_1})$	$\frac{K_{\text{app}}(\mathbf{e_2}) + K_{\text{app}}(\mathbf{e_3})}{2}$
[91] ($K_\perp$) [106, 60]	idem	$\int_0^{2\pi} K_{\text{app}}(\mathbf{e}_\phi)\, d\phi$

Some of these measures have been probed for in vivo and ex vivo rat brain DKI, and compared to their DTI analogues [56]. Mean and radial kurtoses showed strong contrast between GM and WM both in and ex vivo. In particular, radial kurtosis performs better than all other directional diffusivities and kurtoses. For axial kurtosis, a stronger contrast was observed under ex vivo conditions. On the other hand, kurtosis anisotropy was similar to FA both in and ex vivo.

Mean kurtosis and kurtosis anisotropy have also been computed by an adaptive spherical integral, and compared to those based on D-eigenvalues for real diffusion data of a healthy subject and a stroke patient [81]. The latter are seen to be more sensitive to noise. Exact expressions for mean and radial kurtoses can be obtained [106]. These have been shown, together with axial kurtosis, on DKI scans of healthy subjects [106, 60]. The optimization of the diffusion gradient settings for estimation of mean and radial kurtosis, and kurtosis anisotropy has been studied as well. It has been shown that this increases precision considerably [91].

D-eigenvalues of the fourth-order kurtosis tensor $\mathscr{W}$ are defined by [95]

$$\mathscr{W} \cdot^3 \mathbf{x} = \beta\, \mathbf{Dx}; \quad \mathbf{x}^{\mathrm{T}}\mathbf{Dx} = 1, \tag{39}$$

where $\mathbf{x}$ is the D-eigenvector associated to D-eigenvalue β. D-eigenvalues have been shown to be rotationally invariant [95]. The largest and smallest D-eigenvalues can be used to compute the largest and smallest AKC values as $(\text{MD})^2\beta_{\max}$ and $(\text{MD})^2\beta_{\min}$. Other type of eigenvalues which have been studied in this context are the Kelvin eigenvalues of the kurtosis tensor, which are also rotationally invariant. A three-dimensional symmetric fourth-order tensor can be mapped to a six-dimensional second-order tensor. The eigenvalues $(\eta_1, \ldots, \eta_6)$ of its matrix representation, a symmetric 6×6 matrix, are the Kelvin eigenvalues of the considered fourth-order tensor. It has been shown that the largest and smallest Kelvin eigenvalues of (a scaled version of) the kurtosis tensor $\hat{\mathscr{W}}$ are, respectively, an upper and lower bound of the largest and smallest AKC values [94]. The interpretation of Kelvin eigenvalues in terms of AKC values is thus less clear than for D-eigenvalues.

5.1.3 Orientation Distribution Functions

ODF maxima are characterized by their position and value (see Section 5.3), but also by their geometric shape. A peak sharpness measure

$$\mathrm{PS} = \frac{-\mu_1}{k F(\mathbf{u})} \tag{40}$$

can be derived from the value $F(\mathbf{u})$, order k of $\mathscr{F}$, and a Hessian eigenvalue μ_1 of F (at maxima, $\mu_2 \leq \mu_1 \leq 0$). The homogeneous forms of second-order tensors $\mathbf{F}$ have a single maximum, whose sharpness depends on the degree to which $\mathbf{F}$ has a linear shape, as measured by the widely used invariant $c_l = (\lambda_1 - \lambda_2)/\lambda_1$ [115]. In fact, when applied to a second-order tensor, $\mathrm{PS} = c_l$ [99].

Peak Fractional Anisotropy (PFA) is designed to coincide with traditional Fractional Anisotropy (FA) [21] when the diffusion process is well-described by a second-order diffusion tensor, but generalizes it to a per-peak measure in case of more than one ODF maximum [41]. It is defined by fitting a second-order tensor to each ODF peak and computing its FA. Based on the function value F and principal curvatures $\kappa_1 > \kappa_2$ at the maximum, the fitted tensor eigenvalues are given by:

$$\mathrm{ODF\text{-}T}: \quad \lambda_1 = F^2, \; \lambda_2 = \tfrac{F}{\kappa_2}, \; \lambda_3 = \tfrac{F}{\kappa_1} \tag{41}$$

$$\mathrm{ODF\text{-}SA}: \lambda_1 = 1, \; \lambda_2 = \tfrac{3}{2+\kappa_2 F}, \; \lambda_3 = \tfrac{3}{2+\kappa_1 F} \tag{42}$$

ODF-T refers to the q-ball defined by Tuch [110]; ODF-SA denotes a solid angle ODF [1, 109]. The total PFA is defined by considering a weighted sum of the PFA over all ODF maxima:

$$\mathrm{Total\text{-}PFA} = \sum_{i=1}^{\sharp\mathrm{max.}} F_i \cdot \mathrm{PFA}_i \tag{43}$$

Unlike Fractional Anisotropy, Total-PFA is able to distinguish between near-isotropic regions with many weak ODF maxima and areas with complex fiber structure, which exhibit multiple, high anisotropy maxima.

Other geometrical scalars have also been considered. The Ricci scalar is a well-known invariant quantity in differential geometry representing intrinsic curvature, and constructed from the metric and metric-derived tensors. It has been proposed as a DTI scalar measure in the context of Riemannian geometry [35]. The Ricci scalar can also be calculated from a (strongly) convexified ODF by relating it to Finsler geometry (see Section 5.5 and Chapter **??**) [6]. However, experimental results on the latter have not yet been reported.

In addition, principal invariants of fully symmetric fourth-order tensors representing an ODF have been studied [36]. Invariants of fourth-order covariance tensors in DTI had been previously investigated [20]. More general invariants of fourth-order tensors have been recently presented [42]. Principal invariants can be computed from the tensor Kelvin eigenvalues $(\eta_1, \ldots, \eta_6)$ (see section 5.1.2):

$$I_1 = \eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5 + \eta_6$$
$$I_2 = \eta_1\eta_2 + \eta_1\eta_3 + \ldots + \eta_5\eta_6$$
$$I_3 = \eta_1\eta_2\eta_3 + \eta_1\eta_2\eta_4 + \ldots + \eta_4\eta_5\eta_6 \qquad (44)$$
$$I_4 = \eta_1\eta_2\eta_3\eta_4 + \eta_1\eta_2\eta_3\eta_5 + \ldots + \eta_3\eta_4\eta_5\eta_6$$
$$I_5 = \eta_1\eta_2\eta_3\eta_4\eta_5 + \ldots + \eta_2\eta_3\eta_4\eta_5\eta_6$$
$$I_6 = \eta_1\eta_2\eta_3\eta_4\eta_5\eta_6$$

These quantities are, by definition, rotationally invariant and can therefore be used as building blocks for invariant scalar HARDI measures. Experiments on HARDI phantom and brain data have been presented but further work is required to asses the utility of principal invariants in this context. Finally, note that both the Ricci scalar and principal invariants can also be calculated from higher-order diffusion tensors.

5.2 Reconstructing the Diffusion Propagator

The diffusion process is characterized by a probability density function $P(\mathbf{r}, t)$ that specifies the probability of a spin displacement $\mathbf{r}$ within diffusion time t. $P(\mathbf{r}, t)$ is known as the diffusion propagator or Ensemble Average Propagator (EAP). It is related to the dMRI signal by a Fourier transform in the q-space formalism $S(\mathbf{q}, t)/S_0 = \int_{\mathbb{R}^3} P(\mathbf{r}, t) e^{2\pi i \mathbf{q} \cdot \mathbf{r}} d\mathbf{r}$ [25]. Even though higher order tensor estimates of ADC and kurtosis can discern regions with multiple fiber directions, they cannot be used to resolve the directions themselves. To resolve fiber directions the EAP or its characteristics such as the ODF need to be computed.

In DTI, the diffusivities are modelled by a quadratic function given by the diffusion tensor, Eq. (1). The Fourier transform of the resulting signal yields the corresponding EAP, an oriented Gaussian distribution with the tensor's largest eigen-pair indicating the single major fiber direction. However, when HOTs are used to model more complex ADC profiles, computing the EAP turns out to be a trickier problem.

Unlike in DTI, the analytical Fourier transform of the tensor model in Eq. (2) is unknown. In [88], a fast Fourier transform was performed on interpolated (and extrapolated) q-space data on a Cartesian grid generated from the tensor in Eq. (2) to numerically estimate the EAP. In [87], an analytical EAP on a single R_0-shell, i.e., $P(R_0 \frac{\mathbf{r}}{\|\mathbf{r}\|})$, was proposed for this model. However, in this Diffusion Orientation Transform (DOT), the SH basis representation of the tensor was used, see Eq. (21).

In [40], the authors considered a modified non-monoexponential model inspired from Eq. (2) where the HOT was used to describe the signal in the entire q-space. The modified model leads to an analytical series expansion of the EAP in Hermite polynomials. In [15], the authors proposed to use tensors to describe a single R_0-shell of the EAP, $P(R_0 \frac{\mathbf{r}}{\|\mathbf{r}\|})$. They used Hermite polynomials to describe the dMRI signal, since under certain constraints the Fourier transform of Hermite polynomials are homogeneous forms or tensors. Note that [40] and [15] used the same dual Fourier bases but in the opposite spaces to analytically resolve the Fourier transform.

The first attempt to estimate the EAP analytically was based on the tensor model in Eq. (3), where the HOTs represented the cumulant tensors of the EAP since the dMRI signal is also the characteristic function of the EAP. The authors in [78, 77], proposed to use the Gram-Charlier series to compute a series estimate of the EAP from the first four cumulant tensors, i.e. covariance (diffusion) and kurtosis. In theory, the Gram-Charlier series could be improved by the Edgeworth series [45].

In [69], the authors computed the ODF directly from the first four cumulant tensors – diffusion and kurtosis. In contrast to [78, 77], they do not estimate the full EAP, but only its radial marginalization.

5.3 Finding Maxima of the Homogeneous Form

The maxima of many orientation distribution functions in dMRI, which can be represented in the HOT or SH bases, indicate underlying fiber directions. It is, therefore, crucial to compute these maxima with high precision.

The simplest approach is to discretely sample the homogeneous form on a spherical mesh and to compare its values on the finite vertices to approximately identify the maxima [54]. However, even a 16th order tessellation of the icosahedron or 1281 vertices on the sphere can lead to an error of $\sim 4°$. Numerical optimization techniques such as Newton-Raphson and Powell's methods have been used in the SH basis [58, 108] to overcome this limitation. In [55], numerical optimization was combined with the Euler integration step of a tractography algorithm in the tensor basis to trace fibers efficiently.

However, such local optimization techniques are highly dependent on initialization. In [23] and [47] two methods were shown for computing all the stationary points of a homogeneous form. In [23], the Z-eigenvalue/eigenvector formulation was used and a system of two polynomials in two variables – the homogeneous form and the unit sphere constraint – was solved using resultants (detailed in [96]). The stationary points were then classified by their principal curvatures into maxima, minima and saddle-points. In [46], the gradient of the homogeneous form constrained to the unit sphere – a system of four polynomials – was equated to zero. The roots of the system were computed by the subdivision method which ensures that all roots are analytically bracketed thus missing none. The stationary points were then classified into maxima, minima and saddle points using the Bordered Hessian.

5.4 Tensor Decompositions and Approximations

There are four lines of work that have applied tensor decompositions in the context of diffusion MRI. The first results from considering normal distributions of second-order diffusion tensors, which involve a fourth-order covariance tensor Σ. When the diffusion tensor is written as a vector, Σ is naturally represented by a 6×6 symmet-

ric positive definite matrix **S**, to which the spectral decomposition into eigenvalues and eigentensors can be applied, in order to facilitate visualization and quantitative analysis [20]. Alternatively, Σ can be expressed in a local coordinate frame that is derived from invariant gradients and rotation tangents [63]. The coordinates in this frame isolate physically and biologically meaningful components such as variability that can be attributed to changes in trace, anisotropy, or orientation.

Second, the distribution of fiber orientation estimates, either from the diffusion tensor or from HARDI, has been modeled by mapping the corresponding probability measure into a reproducing kernel Hilbert space. With a power-of-cosine kernel, this results in a higher-order tensor representation, which can be decomposed into a rank-one approximation and a non-negative residual to visually and quantitatively investigate the uncertainty in fiber estimates from diffusion MRI [100].

Third, in the framework described in detail in Section 4.3, a low-rank approximation of fODF tensors provides a less biased estimate of principal directions than fODF maxima. It has been shown [102] that this model can be used to approximate and to more efficiently and robustly fit the ball-and-multi-stick model [22]. Subsequent work has imposed an additional non-negativity constraint during deconvolution, and proposed an alternative optimization algorithm [62]. Low-rank approximations were shown to produce useful estimates of crossing fibers even from a relatively small number of gradient directions [49].

Finally, another line of work has attempted to decompose higher-order diffusion tensors in order to obtain crossing fiber directions [116, 59]. However, these techniques are yet to be validated on synthetic data with varying crossing angles, and have not yet been shown to reconstruct known fiber crossings in real data.

5.5 *Finslerian Tractography*

DTI streamline tracking can be generalized to HARDI by means of Finsler geometry. A second-order Finsler metric tensor can be defined at each point **q** from an ODF in the following way [4, 34, 7, 5]

$$\hat{F}(\mathbf{q},\mathbf{x}) = \left(\sum_{i_i \dots i_p} \mathscr{F}_{i_1 \dots i_p}(\mathbf{q}) x^{i_1} \dots x^{i_p}\right)^{1/p}, \ \ g_{ij}(\mathbf{q},\mathbf{x}) = \frac{1}{2}\frac{\partial \hat{F}^2(\mathbf{q},\mathbf{x})}{\partial x^i \partial x^j} \qquad (45)$$

where $\mathscr{F}$ is an ODF tensor of (even) order p, $\hat{F}$ is the Finsler function and g_{ij} is the Finsler metric, $i,j = 1,\dots,3$, which depends on both position and direction. Note that this definition of the Finsler function $\hat{F}$ is by no means unique. In fact, this is still a subject of intensive research (see Chapter **??**). Thus a local diffusion tensor can be obtained per direction. Tracking can be performed by extracting the principal eigenvector of the diffusion tensor corresponding to the arrival direction. As long as this direction is sufficiently aligned to the eigenvector, and the diffusion tensor FA is above a certain treshold, tracking continues. Experiments on Finsler streamline tracking using fourth-order tensors have been presented on simulated fiber crossings

and real HARDI data. It has been shown that Finsler streamlines can, unlike DTI streamlines, correctly cope with nerve fiber bundle crossings.

5.6 Registration and Atlas Construction

Registration transforms data sets from different times or subjects to a common coordinate system, so that anatomical structures align. Atlas construction is based on registering a large number of subjects, in order to obtain a description of average anatomy, and of the most common modes of variation. Modeling parameters of the diffusion process with higher-order tensors makes registration of tensor fields a relevant research problem. Registration requires selection of an appropriate metric to measure the dissimilarity between individual tensors; for this purpose, Barmpoutis et al. [14, 12] propose two alternative choices, which are both scale and rotation invariant. Integrating the local dissimilarity over the domain of the tensor field results in an overall measure of dissimilarity. Registration is achieved by finding the coordinate transformation that minimizes this measure.

It is important to also transform the individual tensors according to the coordinate transformation applied to the domain of the field. For example, when the domain of the tensor field is rotated, a corresponding rotation of the tensors themselves is required in order to preserve relevant structures, such as the trajectories of nerve fiber bundles. When the transformation is (locally) affine, it has been proposed to simply apply it to the tensors via Eq. (11) [14]. Alternative methods for transformation have been proposed based on the spectral decomposition [97] and different sum-of-squares parametrizations [9, 97, 48].

6 Conclusion

The wide range of models and computational methods that have been surveyed in this chapter testify to the power and flexibility that higher-order tensors provide for the analysis of data from diffusion MRI, and to the increasing momentum of the research associated with this topic. Generalized eigenvalues, scalar invariants, tensor decompositions, and low-rank approximations have all proven valuable in the context of this application.

Looking ahead, several *theoretical problems* remain to be solved. While many approaches have focused on the properties of individual tensors, less attention has been paid to the *global nature of the tensor fields* that arise in diffusion MRI. The recent use of Finsler geometry is a natural step in this direction.

Even though *low-rank approximations* have proven to work well in practice, uniqueness of approximations over the reals is mostly open (for the complex case, see [66]). Moreover, we are still lacking algorithms with provable convergence properties, and formal results on the well-conditionedness of such approximations.

Many approaches have been proposed to ensure *non-negativity of higher-order tensors* that model apparent diffusivities (cf. Section 4.1). Less attention has been paid to the fitting of deconvolution models, which are constrained to the convex cone of tensors that can be expressed as a positive sum of rank-1 tensors; in general, that is a stricter constraint than non-negativity.

While many neuroscientific studies that use diffusion imaging are now published each month, they still almost exclusively use either the second-order diffusion tensor [21] or the ball-and-stick model [22]. A challenge in the next few years will be to *take approaches based on higher-order tensors into the application domain.* This will require more work on several subproblems:

Statistical tests on *scalar invariants* such as Mean Diffusivity or Fractional Anisotropy are a mainstay of DTI-based studies. Even though a considerable number of invariants have now been derived from higher-order tensors (cf. Section 5.1), the practical utility of many of them is limited by their unclear biological or neuroanatomical interpretation.

Given an ever-increasing palette of models, it becomes a more urgent problem to pick one of them to test a given hypothesis, and to choose values for parameters such as tensor order, approximation rank, or regularization weights. Improved understanding of *formal relationships between different models* and mathematical rules for *model selection* are required.

Spatial coherence and *signal sparsity* need to be exploited in order to reliably estimate the large number of parameters in advanced models such as the ensemble average propagator, without requiring excessively time consuming measurements.

Acknowledgments

References

1. Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., Harel, N.: Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. Magnetic Resonance in Medicine **64**(2), 554–566 (2010)
2. Anderson, A.W.: Measurement of fiber orientation distributions using high angular resolution diffusion imaging. Magnetic Resonance in Medicine **54**(5), 1194–1206 (2005)
3. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-euclidean metrics for fast and simple calculus on diffusion tensors. Magnetic Resonance in Medicine **56**(2), 411–421 (2006)

4. Astola, L., Florack, L.: Finsler geometry on higher order tensor fields and applications to high angular resolution diffusion imaging. In: Proc. Int'l Conf. on Scale Space and Variational Methods in Computer Vision (SSVM), pp. 224–234. Springer (2009)

5. Astola, L., Florack, L.: Finsler geometry on higher order tensor fields and applications to high angular resolution diffusion imaging. Int'l J. of Computer Vision **92**, 325–336 (2011)

6. Astola, L., Fuster, A., Florack, L.: A Riemannian scalar measure for diffusion tensor images. Pattern Recognition **44**, 1885–1891 (2011)

7. Astola, L., Jalba, A., Balmashnova, E., Florack, L.: Finsler streamline tracking with single tensor orientation distribution function for high angular resolution diffusion imaging. J. of Mathematical Imaging and Vision **41**, 170–181 (2011)

8. Balmashnova, E., Fuster, A., Florack, L.: Decomposition of higher-order homogeneous tensors and applications to HARDI. In: E. Panagiotaki, L. O'Donnell, T. Schultz, G.H. Zhang (eds.) Proc. Computational Diffusion MRI (CDMRI), pp. 79–89 (2012)

9. Barmpoutis, A., Hwang, M.S., Howland, D., Forder, J.R., Vemuri, B.C.: Regularized positive-definite fourth order tensor field estimation from DW-MRI. NeuroImage **45**(1, Supl. 1), S153–S162 (2009)

10. Barmpoutis, A., Jian, B., Vemuri, B.C., Shepherd, T.M.: Symmetric positive 4th order tensors & their estimation from diffusion weighted MRI. In: N. Karssemeijer, B. Lelieveldt (eds.) IPMI, *LNCS*, vol. 4584, pp. 308–319 (2007)

11. Barmpoutis, A., Vemuri, B.C.: Exponential tensors: A framework for efficient higher-order DT-MRI computations. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 792–795 (2007)

12. Barmpoutis, A., Vemuri, B.C.: Groupwise registration and atlas construction of 4th-order tensor fields using the $\mathbb{R}^+$ Riemannian metric. In: G.Z. Yang, D. Hawkes, D. Rueckert, A. Noble, C. Taylor (eds.) Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), Part I, *LNCS*, vol. 5761, pp. 640–647 (2009)

13. Barmpoutis, A., Vemuri, B.C.: A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 1385–1388 (2010)

14. Barmpoutis, A., Vemuri, B.C., Forder, J.R.: Registration of high angular resolution diffusion MRI images using 4th order tensors. In: N. Ayache, S. Ourselin, A. Maeder (eds.) Proc. Medical Image Computing and Computer-Assisted Intervention, Part I, *LNCS*, vol. 4791, pp. 908–915 (2007)

15. Barmpoutis, A., Vemuri, B.C., Forder, J.R.: Fast displacement probability profile approximation from HARDI using 4th-order tensors. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 911–914 (2008)

16. Barmpoutis, A., Zhuo, J.: Diffusion kurtosis imaging: Robust estimation from DW-MRI using homogeneous polynomials. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 262–265 (2011)

17. Barnett, A.: Theory of Q-ball imaging redux: Implications for fiber tracking. Magnetic Resonance in Medicine **62**(4), 910–923 (2009)

18. Basser, P.J., Mattiello, J., Le Bihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. Journal of Magnetic Resonance **B**(103), 247–254 (1994)

19. Basser, P.J., Pajevic, S.: A normal distribution for tensor-valued random variables: Applications to diffusion tensor MRI. IEEE Trans. on Medical Imaging **22**, 785–795 (2003)

20. Basser, P.J., Pajevic, S.: Spectral decomposition of a 4th-order covariance tensor: Applications to diffusion tensor MRI. Signal Processing **87**, 220–236 (2007)

21. Basser, P.J., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. Journal of Magnetic Resonance, Series B **111**, 209–219 (1996)

22. Behrens, T.E.J., Johansen-Berg, H., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage **34**, 144–155 (2007)

23. Bloy, L., Verma, R.: On computing the underlying fiber directions from the diffusion orientation distribution function. In: D.N. Metaxas, L. Axel, G. Fichtinger, G. Székely (eds.) Medical Image Computing and Computer-Assisted Intervention (MICCAI), *LNCS*, vol. 5241, pp. 1–8. Springer (2008)

24. Boothby, W.: An Introduction to Differentiable Manifolds and Riemannian Geometry, *Pure and Applied Mathematics*, vol. 120, 2nd edn. Academic Press, Orlando, FL (1986)

25. Callaghan, P.T., Eccles, C.D., Xia, Y.: NMR microscopy of dynamic displacements: k-space and q-space imaging. Journal of Physics E **21**(8), 820–822 (1988)

26. Comon, P., Golub, G., Lim, L.H., Mourrain, B.: Symmetric tensors and symmetric tensor rank. SIAM Journal on Matrix Analysis and Applications **30**(3), 1254–1279 (2008)

27. Correia, M.M., Newcombe, V.F., Williams, G.B.: Contrast-to-noise ratios for indices of anisotropy obtained from diffusion MRI: A study with standard clinical b-values at 3T. NeuroImage **57**(3), 1103–1115 (2011)

28. De Silva, V., Lim, L.H.: Tensor rank and the ill-posedness of the best low-rank approximation problem. SIAM Journal on Matrix Analysis and Applications **30**(3), 1084–1127 (2008)

29. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. Magnetic Resonance in Medicine **56**, 395–410 (2006)

30. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-Ball imaging. Magnetic Resonance in Medicine **58**, 497–510 (2007)

31. Descoteaux, M., Deriche, R., Knösche, T.R., Anwander, A.: Deterministic and probabilistic tractography based on complex fibre orientation distributions. IEEE Transactions on Medical Imaging **28**(2), 269–286 (2009)

32. Ellingson, B.M., Cloughesy, T.F., Lai, A., Nghiemphu, P.L., Liau, L.M., Pope, W.B.: High order diffusion tensor imaging in human glioblastoma. Academic Radiology **18**(8), 947–954 (2011)

33. Essen, D.V., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S., Penna, S.D., Feinberg, D., Glasser, M., Harel, N., Heath, A., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S., Prior, F., Schlaggar, B., Smith, S., Snyder, A., Xu, J., Yacoub, E.: The human connectome project: A data acquisition perspective. NeuroImage **62**(4), 2222–2231 (2012)

34. Florack, L., Balmashnova, E., Astola, L., Brunenberg, E.: A new tensorial framework for single-shell high angular resoltion diffusion imaging. J. of Mathematical Imaging and Vision **38**, 171–181 (2010)

35. Fuster, A., Astola, L., Florack, L.: A Riemannian scalar measure for diffusion tensor images. In: X. Jiang, N. Petkov (eds.) Computer Analysis of Images and Patterns, *LNCS*, vol. 5702, pp. 419–426. Springer (2009)

36. Fuster, A., van de Sande, J., Astola, L., Poupon, C., Velterop, J., ter Haar Romeny, B.M.: Fourth-order tensor invariants in high angular resolution diffusion imaging. In: G.H. Zhang, N. Adluru (eds.) Proc. MICCAI Workshop on Computational Diffusion MRI, pp. 54–63 (2011)

37. Gelfand, I., Kapranov, M., Zelevinsky, A.: Discriminants, Resultants, and Multidimensional Determinants. Birkhäuser Publishing, Boston, MA (1994)

38. Geroch, R.: Mathematical Physics. Chicago Lectures in Physics. University of Chicago Press, Chicago, IL (1985)

39. Ghosh, A., Deriche, R.: From second to higher order tensors in diffusion-MRI. In: S. Aja-Fernández, R. de Luis García, D. Tao, X. Li (eds.) Tensors in Image Processing and Computer Vision, pp. 315–334. Springer (2009)

40. Ghosh, A., Deriche, R.: Fast and closed-form ensemble-average-propagator approximation from the 4th-order diffusion tensor. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 1105–1108 (2010)

41. Ghosh, A., Deriche, R.: Extracting geometrical features & peak fractional anisotropy from the ODF for white matter characterization. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 266–271 (2011)

42. Ghosh, A., Deriche, R.: Generalized invariants of a 4th order tensor: Building blocks for new biomarkers in dMRI. In: E. Panagiotaki, L. O'Donnell, T. Schultz, G.H. Zhang (eds.) Proc. Computational Diffusion MRI (CDMRI), pp. 165–173 (2012)

43. Ghosh, A., Deriche, R., Moakher, M.: Ternary quartic approach for positive 4th order diffusion tensors revisited. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 618–621 (2009)

44. Ghosh, A., Descoteaux, M., Deriche, R.: Riemannian framework for estimating symmetric positive definite 4th order diffusion tensors. In: D. Metaxas (ed.) MICCAI, Part I, *LNCS*, vol. 5241, pp. 858–865 (2008)

45. Ghosh, A., Özarslan, E., Deriche, R.: Challenges in reconstructing the propagator via a cumulant expansion of the one-dimensional qspace MR signal. In: Proc. Int'l. Soc. for Magnetic Resonance in Medicine (ISMRM) (2010)

46. Ghosh, A., Tsigaridas, E., Mourrain, B., Deriche, R.: A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI. Medical Image Analysis (2013). In press. DOI: 10.1016/j.media.2013.03.004

47. Ghosh, A., Wassermann, D., Deriche, R.: A polynomial approach for maxima extraction and its application to tractography in HARDI. In: G. Székely, H.K. Hahn (eds.) IPMI, *LNCS*, vol. 6801, pp. 723–734 (2011)

48. Grigis, A., Renard, F., Noblet, V., Heinrich, C., Heitz, F., Armspach, J.P.: A new high order tensor decomposition: Application to reorientation. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 258–261 (2011)

49. Gur, Y., Jiao, F., Zhu, S.X., Johnson, C.R.: White matter structure assessment from reduced HARDI data using low-rank polynomial approximations. In: E. Panagiotaki, L. O'Donnell, T. Schultz, G.H. Zhang (eds.) Proc. Computational Diffusion MRI (CDMRI), pp. 186–197 (2012)

50. Hess, C.P., Mukherjee, P., Han, E.T., Xu, D., Vigneron, D.B.: Q-Ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. Magnetic Resonance in Medicine **56**, 104–117 (2006)

51. Hillar, C., Lim, L.H.: Most tensor problems are NP-hard (2012). Preprint, arXiv:0911.1393v2

52. Hitchcock, F.L.: The expression of a tensor or a polyadic as a sum of products. Journal of Mathematics and Physics **6**(1), 164–189 (1927)

53. Hitchcock, F.L.: Multiple invariants and generalized rank of a p-way matrix or tensor. Journal of Mathematics and Physics **7**(1), 39–79 (1927)

54. Hlawitschka, M., Scheuermann, G.: HOT-lines: Tracking lines in higher order tensor fields. In: C. Silva, E. Gröller, H. Rushmeier (eds.) Proc. IEEE Visualization, pp. 27–34 (2005)

55. Hlawitschka, M., Scheuermann, G., Anwander, A., Knösche, T., Tittgemeyer, M., Hamann, B.: Tensor lines in tensor fields of arbitrary order. In: G. Bebis et al. (ed.) Advances in Visual Computing, *LNCS*, vol. 4841, pp. 341–350. Springer (2007)

56. Hui, E.S., Cheung, M.M., Qi, L., Wu, E.X.: Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis. NeuroImage **42**, 122–134 (2008)

57. Hungerford, T.: Algebra, *Graduate Texts in Mathematics*, vol. 73. Springer, New York, NY (1980)

58. Jansons, K.M., Alexander, D.C.: Persistent angular structure: new insights from diffusion magnetic resonance imaging data. Inverse Problems **19**, 1031–1046 (2003)

59. Jayachandra, M.R., Rehbein, N., Herweh, C., Heiland, S.: Fiber tracking of human brain using fourth-order tensor and high angular resolution diffusion imaging. Magnetic Resonance in Medicine **60**(5), 1207–1217 (2008)

60. Jensen, J.H., Helpern, J.A.: MRI quantification of non-gaussian water diffusion by kurtosis analysis. NMR in Biomedicine **23**(7), 698–710 (2010)

61. Jensen, J.H., Helpern, J.A., Ramani, A., Lu, H., Kaczynski, K.: Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magnetic Resonance in Medicine **53**, 1432–1440 (2005)

62. Jiao, F., Gur, Y., Johnson, C.R., Joshi, S.: Detection of crossing white matter fibers with high-order tensors and rank-k decompositions. In: G. Székely, H.K. Hahn (eds.) IPMI, *LNCS*, vol. 6801, pp. 538–549 (2011)

63. Kindlmann, G., Ennis, D., Whitaker, R., Westin, C.F.: Diffusion tensor analysis with invariant gradients and rotation tangents. IEEE Transactions on Medical Imaging **26**(11), 1483–1499 (2007)

64. Kroonenberg, P.: Applied Multiway Data Analysis. John Wiley, Hoboken, NJ (2008)

65. Kuder, T.A., Stieltjes, B., Bachert, P., Semmler, W., Laun, F.B.: Advanced fit of the diffusion kurtosis tensor by directional weighting and regularization. Magnetic Resonance in Medicine **67**(5), 1401–1411 (2012)

66. Landsberg, J.M.: Tensors: Geometry and Applications, *Graduate Studies in Mathematics*, vol. 128. American Mathematical Society (2012)

67. Lang, S.: Differential and Riemannian Manifolds, *Graduate Texts in Mathematics*, vol. 160, 3rd edn. Springer, New York, NY (1995)

68. Lang, S.: Algebra, *Graduate Texts in Mathematics*, vol. 211, rev. 3rd edn. Springer, New York, NY (2002)

69. Lazar, M., Jensen, J.H., Xuan, L., Helpern, J.A.: Estimation of the orientation distribution function from diffusional kurtosis imaging. Magnetic Resonance in Medicine **60**, 774–781 (2008)

70. Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., Laval-Jeantet, M.: MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. Radiology **161**(2), 401–407 (1986)

71. Lenglet, C., Rousson, M., Deriche, R., Faugeras, O.: Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. J. of Mathematical Imaging and Vision **25**, 423–444 (2006)

72. Lim, L.H.: Singular values and eigenvalues of tensors: a variational approach. In: Proc. IEEE Int'l Workshop on Comput. Adv. Multi-Sensor Adaptive Process. (CAMSAP), pp. 129–132 (2005)

73. Lim, L.H.: Tensors and hypermatrices. In: L. Hogben (ed.) Handbook of Linear Algebra, 2nd edn. CRC Press, Boca Raton, FL (2013)

74. Lim, L.H., Comon, P.: Nonnegative approximations of nonnegative tensors. J. Chemometrics **23**(7–8), 432–441 (2009)

75. Lim, L.H., Comon, P.: Multisensor array processing: Tensor decomposition meets compressed sensing. C. R. Acad. Sci. Paris **338**(6), 311–320 (2010)

76. Lim, L.H., Schultz, T.: Moment tensors and high angular resolution diffusion imaging (2013). Preprint

77. Liu, C., Bammer, R., Acar, B., Moseley, M.E.: Characterizing non-gaussian diffusion by using generalized diffusion tensors. Magnetic Resonance in Medicine **51**(5), 924–937 (2004)

78. Liu, C., Bammer, R., Moseley, M.E.: Generalized diffusion tensor imaging (GDTI): A method for characterizing and imaging diffusion anisotropy caused by non-gaussian diffusion. Israel Journal of Chemistry **43**(1–2), 145–154 (2003)

79. Liu, C., Bammer, R., Moseley, M.E.: Limitations of apparent diffusion coefficient-based models in characterizing non-gaussian diffusion. Magnetic Resonance in Medicine **54**, 419–428 (2005)

80. Liu, C., Mang, S.C., Moseley, M.E.: In vivo generalized diffusion tensor imaging (GDTI) using higher-order tensors (HOT). Magnetic Resonance in Medicine **63**, 243–252 (2010)

81. Liu, Y., Chen, L., Yu, Y.: Diffusion kurtosis imaging based on adaptive spherical integral. IEEE Signal Process. Lett. **18**(4), 243–246 (2011)

82. Lu, H., Jensen, J.H., Ramani, A., Helpern, J.A.: Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. NMR in Biomedicine **19**, 236–247 (2006)

83. Minati, L., Aquino, D., Rampoldi, S., Papa, S., Grisoli, M., Bruzzone, M.G., Maccagnano, E.: Biexponential and diffusional kurtosis imaging, and generalised diffusion-tensor imaging (GDTI) with rank-4 tensors: a study in a group of healthy subjects. Magnetic Resonance Materials in Physics, Biology and Medicine **20**, 241–253 (2007)

84. Minati, L., Banasik, T., Brzezinski, J., Mandelli, M.L., Bizzi, A., Bruzzone, M.G., Konopka, M., Jasinski, A.: Elevating tensor rank increases anisotropy in brain areas associated with intra-voxel orientational heterogeneity (IVOH): a generalised DTI (GDTI) study. NMR in Biomedicine **21**(1) (2008)

85. Mørup, M., Hansen, L., Arnfred, S., Lim, L.H., Madsen, K.: Shift invariant multilinear decomposition of neuroimaging data. NeuroImage **42**(4), 1439–1450 (2008)

86. Özarslan, E., Mareci, T.: Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. Magnetic Resonance in Medicine **50**, 955–965 (2003)

87. Özarslan, E., Shepherd, T.M., Vemuri, B.C., Blackband, S.J., Mareci, T.H.: Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). NeuroImage **31**, 1086–1103 (2006)

88. Özarslan, E., Vemuri, B.C., Mareci, T.H.: Fiber orientation mapping using generalized diffusion tensor imaging. In: Proc. IEEE International Symposium on Biomedical Imaging, pp. 1036–1039 (2004)

89. Özarslan, E., Vemuri, B.C., Mareci, T.H.: Generalized scalar measures for diffusion MRI using trace, variance, and entropy. Magnetic Resonance in Medicine **53**, 866–876 (2005)

90. Özarslan, E., Vemuri, B.C., Mareci, T.H.: Higher rank tensors in diffusion MRI. In: J. Weickert, H. Hagen (eds.) Visualization and Processing of Tensor Fields, chap. 10, pp. 177–187. Springer (2006)

91. Poot, D.H.J., den Dekker, A.J., Achten, E., Verhoye, M., Sijbers, J.: Optimal experimental design for diffusion kurtosis imaging. IEEE Trans. on Medical Imaging **29**(3), 819–829 (2010)

92. Procesi, C.: Lie Groups: An Approach through Invariants and Representation. Springer, New York, NY (2008)

93. Qi, L.: Eigenvalues of a real supersymmetric tensor. Journal of Symbolic Computation **40**, 1302–1324 (2005)

94. Qi, L., Han, D., Wu, E.X.: Principal invariants and inherent parameters of diffusion kurtosis tensors. J. of Mathematical Analysis and Applications **349**, 165–180 (2009)

95. Qi, L., Wang, Y., Wu, E.X.: D-eigenvalues of diffusion kurtosis tensors. J. of Computational and Applied Mathematics **221**, 150–157 (2008)

96. Qi, L., Yu, G., Wu, E.X.: Higher order positive semidefinite diffusion tensor imaging. SIAM J. Imaging Sciences **3**(3), 416–433 (2010)

97. Renard, F., Noblet, V., Heinrich, C., Kremer, S.: Reorientation strategies for high order tensors. In: Proc. IEEE Int'l Symposium on Biomedical Imaging, pp. 1185–1188 (2010)

98. Schultz, T.: Learning a reliable estimate of the number of fiber directions in diffusion MRI. In: N. Ayache et al. (ed.) Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI) Part III, *LNCS*, vol. 7512, pp. 493–500 (2012)

99. Schultz, T., Kindlmann, G.: A maximum enhancing higher-order tensor glyph. Computer Graphics Forum **29**(3), 1143–1152 (2010)

100. Schultz, T., Schlaffke, L., Schölkopf, B., Schmidt-Wilcke, T.: HiFiVE: a hilbert space embedding of fiber variability estimates for uncertainty modeling and visualization. Computer Graphics Forum **32**(3) (2013)

101. Schultz, T., Seidel, H.P.: Estimating crossing fibers: A tensor decomposition approach. IEEE Trans. on Visualization and Computer Graphics **14**(6), 1635–1642 (2008)

102. Schultz, T., Westin, C.F., Kindlmann, G.: Multi-diffusion-tensor fitting via spherical deconvolution: A unifying framework. In: T. Jiang et al. (ed.) Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), *LNCS*, vol. 6361, pp. 673–680. Springer (2010)

103. Sidiropoulos, N., Bro, R., Giannakis, G.: Parallel factor analysis in sensor array processing. IEEE Trans. Signal Process. **48**(8), 2377–2388 (2000)

104. Smilde, A., Bro, R., Geladi, P.: Multi-way Analysis: Applications in the Chemical Sciences. John Wiley, West Sussex (2004)

105. Struik, D.J.: Schouten, Levi-Civita and the emergence of tensor calculus. In: D. Rowe, J. McCleary (eds.) History of Modern Mathematics, vol. 2, pp. 99–105. Academic Press, Boston, MA (1989)

106. Tabesh, A., Jensen, J.H., Ardekani, B.A., Helpern, J.A.: Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. Magnetic Resonance in Medicine **65**, 823–836 (2011)
107. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. NeuroImage **35**, 1459–1472 (2007)
108. Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage **23**, 1176–1185 (2004)
109. Tristan-Vega, A., Westin, C.F., Aja-Fernandez, S.: A new methodology for the estimation of fiber populations in the white matter of the brain with the funk-radon transform. NeuroImage **49**(2), 1301–1315 (2010)
110. Tuch, D.S.: Q-Ball imaging. Magnetic Resonance in Medicine **52**, 1358–1372 (2004)
111. Vasilescu, M., Terzopoulos, D.: Multilinear image analysis for facial recognition. Proc. Int. Conf. Pattern Recognition (ICPR) **2**, 511–514 (2002)
112. Veraart, J., Van Hecke, W., Sijbers, J.: Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a rician noise model. Magnetic Resonance in Medicine **66**, 678–686 (2011)
113. Warner, F.: Foundations of Differentiable Manifolds and Lie Groups, *Graduate Texts in Mathematics*, vol. 94. Springer, New York-Berlin (1983)
114. Weldeselassie, Y.T., Barmpoutis, A., Atkins, M.S.: Symmetric positive-definite cartesian tensor orientation distribution functions (CT-ODF). In: T. Jiang, N. Navab, J.P.W. Pluim, M.A. Viergever (eds.) Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), *LNCS*, vol. 6361, pp. 582–589 (2010)
115. Westin, C.F., Peled, S., Gudbjartsson, H., Kikinis, R., Jolesz, F.A.: Geometrical diffusion measures for MRI from tensor basis analysis. In: Proc. Int'l. Soc. for Magnetic Resonance in Medicine (ISMRM), p. 1742 (1997)
116. Ying, L., Zou, Y.M., Klemer, D.P., Wang, J.J.: Determination of fiber orientation in MRI diffusion tensor imaging based on higher-order tensor decomposition. In: Proc. Int'l. Conf. IEEE Engineering in Medicine and Biology Society (EMBS), pp. 2065–2068 (2007)
117. Yokonuma, T.: Tensor Spaces and Exterior Algebra, *Translations of Mathematical Monographs*, vol. 108. AMS, Providence, RI (1992)

Index