

TGV for diffusion tensors: A comparison of fidelity functions

Tuomo Valkonen^{*,†} Kristian Bredies[‡] and Florian Knoll[§]

October 24, 2012

Abstract

We study the total generalised variation regularisation of symmetric tensor fields from medical applications, namely diffusion tensor regularisation. We study the effect of the pointwise positivity constraint on the tensor field, as well as the difference between direct denoising of the tensor field first solved from the Stejskal-Tanner equation, as was done in our earlier work, and of incorporating this equation into the fidelity function. Our results indicate that the latter novel approach provides improved computational results.

Mathematics subject classification: 92C55, 94A08, 26B30, 49M29.

Keywords: total generalised variation, total deformation, regularisation, diffusion tensor imaging, medical imaging.

1 Introduction

In this manuscript, we present a numerical study of total generalised variation (TGV) regularisation [6] of second-order symmetric tensor fields $u \in L^1(\Omega; \text{Sym}^2(\mathbb{R}^m))$, namely the problem

$$\min_{u \geq 0} \frac{1}{2} \|f - \mathcal{A}u\|^2 + \text{TGV}_\alpha^\ell(u), \quad (\text{P})$$

where we have a pointwise a.e. positive semi-definiteness constraint on u . The measurement $f \in L^1(\Omega; X)$ for a finite-dimensional Hilbert space X , and the operator \mathcal{A} is the pointwise superposition operator associated with a linear operator $A : \text{Sym}^2(\mathbb{R}^m) \rightarrow X$, where the tensor space $\text{Sym}^2(\mathbb{R}^m)$ consists of symmetric bilinear mappings on $\mathbb{R}^m \times \mathbb{R}^m$. The TGV^ℓ regularisation functionals extended to tensor fields employ the *symmetrised distributional derivative* Eu of u instead of the full derivative Du in order to obtain weaker gradient information. The former, which we have found to produce improved denoising results [23], may for $u \in C^1(\Omega; \text{Sym}^2(\mathbb{R}^m))$ be *formally* written as $Eu = (Du + (Du)^T)/2$. For *symmetric tensor fields of bounded deformation* the distribution Eu is a measure with values in $\text{Sym}^3(\mathbb{R}^m)$, i.e., third-order symmetric tensors. First order TGV, also called *total deformation* or TD, thus reads

$$\text{TGV}_\alpha^1(u) := \alpha \text{TD}(u) := \alpha \|Eu\|_{F, \mathcal{M}(\Omega; \text{Sym}^3(\mathbb{R}^m))}$$

where $\|\cdot\|_{F, \mathcal{M}(\Omega; \text{Sym}^k(\mathbb{R}^m))}$ is the Radon norm. Solutions to first-order regularisation problem tend to exhibit staircasing artefacts, as we will also see in the numerical experiments that we perform in this paper. In order to reduce these effects, we employ second-order *total generalised variation* TGV^2 [6], which for symmetric tensor fields reads

$$\text{TGV}_{(\beta, \alpha)}^2(u) := \min_{w \in L^1(\Omega; \text{Sym}^3(\mathbb{R}^m))} \alpha \|Eu - w\|_{F, \mathcal{M}(\Omega; \text{Sym}^3(\mathbb{R}^m))} + \beta \|Ew\|_{F, \mathcal{M}(\Omega; \text{Sym}^4(\mathbb{R}^m))}.$$

^{*}Institute for Mathematics and Scientific Computing, University of Graz, Austria. tuomo.valkonen@iki.fi

[†]Now at Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom.

[‡]Institute for Mathematics and Scientific Computing, University of Graz, Austria. kristian.bredies@uni-graz.at

[§]Institute of Medical Engineering, Graz University of Technology, Austria. florian.knoll@tugraz.at

The main motivation for considering the problem (P) is diffusion-weighted magnetic resonance imaging (DWI, diffusion MRI) which images a tensor field g describing the anisotropic diffusion of water molecules, providing insight into the white matter structure of the brain [3, 21] with the objective of performing medical diagnosis. Diffusion-weighted MRI has long acquisition times, even with ultra fast sequences like echo planar imaging (EPI). It is therefore a low-resolution and low-SNR method, exhibiting Rician noise [13] and eddy-current distortions [21]. It is, moreover, very sensitive to patient motion [14, 1]. It is therefore of utmost importance to denoise the diffusion tensor field g . Various approaches are reviewed in [23], with the focus on the direct denoising of g by (P) with $A = I$ the identity, $X = \text{Sym}^2(\mathbb{R}^3)$, and $f = g$. As in DWI, the (noisy) tensor field is not measured directly, we have to assume that g is already given as the outcome of some reconstruction process. Under such an assumption, we showed that TGV^2 , first introduced in [6] for scalar fields, as well as *total deformation* (TD), both based on the symmetrised gradient, offer superior denoising performance over other evaluated approaches. These include total variation (TV) [18], and “logarithmic TV” [11], which is based on log-Euclidean metrics [2].

In this paper, we incorporate the above-mentioned reconstruction process into the denoising problem. Let, for $i = 0, \dots, K$ be the varying diffusion gradients $b_i \in \mathbb{R}^3$ associated with the DWI measurement, including the zero gradient $b_0 = 0$. The measured data then consists of $s_i \in L^1(\Omega)$ ($i = 0, \dots, K$), which is related to the sought *diffusion tensor field* $u \in L^1(\Omega; \text{Sym}^2(\mathbb{R}^3))$ via the Stejskal-Tanner equation [3, 15]

$$s_i = s_0 \exp(-\langle b_i \otimes b_i, u \rangle). \quad (1.1)$$

If $b_i \otimes b_i$ form a (necessarily non-orthogonal) basis of $\text{Sym}^2(\mathbb{R}^3)$, we may solve this equation for u . We want to see what effect incorporating (1.1) into (P) has. However, in order to get a linear inversion problem, we reformulate the Stejskal-Tanner equation. Observe that

$$\log(s_i/s_0) + \langle b_i \otimes b_i, u \rangle = 0.$$

Therefore, with $X = \mathbb{R}^K$, setting

$$f = (\log(s_1/s_0), \dots, \log(s_K/s_0)) \quad (1.2)$$

and $(\mathcal{A}u)(x) = Au(x)$ for $x \in \Omega$ where

$$Av = (-\langle b_1 \otimes b_1, v \rangle, \dots, -\langle b_K \otimes b_K, v \rangle), \quad (1.3)$$

we have a problem of the form $\mathcal{A}(u) = f$. As the measurements s_i are noisy, we choose to regularise with $\text{TGV}_\alpha^\ell(u)$ which leads to a problem of type (P). We will compare this against the approach of [23] for both TGV^2 , and TGV^1 , the latter agreeing with total deformation (TD). We moreover study the effect of the pointwise positivity constraint $u \geq 0$.

The solution of the positivity constrained problem (P) poses some numerical challenges. We employ the Chambolle-Pock method [9], which is a first-order method for solving saddle-point problems of the form

$$\min_u \max_\phi G(u) + \langle \mathcal{K}u, \phi \rangle - F^*(\phi). \quad (1.4)$$

With the reformulation of (P) in the form (1.4) that was employed in [23], the computational hotspot of the algorithm is the computation on each iteration of the resolvent $(I + \tau \partial G)^{-1}(y)$ of

$$G(u) := \frac{1}{2} \|f - \mathcal{A}u\|^2 + \delta_{\geq 0}(u), \quad (1.5)$$

where the indicator function $\delta_{\geq 0}$ replaces the pointwise positivity constraint. In the case $A = I$, the calculation of the resolvent amounts to pointwise projection of a second-order tensor to the positive semi-definite cone. This is computationally demanding, as in practise the QR algorithm is required for numerical stability. Nevertheless, the overall method has turned out to be reasonably efficient, minding the non-smooth nature of the problem, which complicates the application of second-order

methods. For the case (1.2), (1.3), we have $A \neq I$, and sticking to the same approach, to compute the resolvent, we have to perform “weighted projections”

$$\min_{0 \leq X \in \text{Sym}^2(\mathbb{R}^m)} \|AX - c\|^2, \quad (1.6)$$

which are more demanding to solve. We have developed a novel simple and efficient interior point method for this projection in [22]. We have, however, experimentally observed that the Chambolle-Pock method itself has poor convergence when A^*A is far from proportional to the identity and G as in (1.5). We therefore introduce an additional dual variable and write (P) in an alternative way in the same form (1.4). Although this formulation has poorer theoretical convergence properties, lacking uniform convexity, and indeed does not provide extremely fast convergence, it appears to provide better convergence than our earlier formulation when A^*A is far from proportional to the identity.

The rest of this paper is organised as follows. In Section 2 we introduce the tensor and tensor field calculus to set up the framework in which our results are represented. Then in Section 3 we formulate the problem (P) in detail. Section 4 describes our numerical method for the solution of (P). Then, in Section 5 we study the performance of our method. We finish the paper with our conclusions in Section 6.

2 Tensors and tensor fields

We now recall basic tensor calculus, as needed for the development of TGV². We make many simplifications, as, working on the Euclidean space \mathbb{R}^m , it would not serve our purposes to employ and introduce the machinery in its full differential-geometric setting [4].

Basic tensor calculus By a k -tensor $A \in \mathcal{T}^k(\mathbb{R}^m)$ we mean a k -linear mapping $A : \mathbb{R}^m \times \cdots \times \mathbb{R}^m \rightarrow \mathbb{R}$. A symmetric tensor $A \in \text{Sym}^k(\mathbb{R}^m)$ satisfies for any permutation π of $\{1, \dots, k\}$ that $A(c_{\pi 1}, \dots, c_{\pi k}) = A(c_1, \dots, c_k)$ for all $c_1, \dots, c_k \in \mathbb{R}^m$. We will only need symmetric tensors in this work. With e_1, \dots, e_m the standard basis of \mathbb{R}^m , we define the inner product

$$\langle A, B \rangle := \sum_{p \in \{1, \dots, m\}^k} A(e_{p_1}, \dots, e_{p_k}) B(e_{p_1}, \dots, e_{p_k}),$$

and the induced Frobenius norm

$$\|A\|_F := \sqrt{\langle A, A \rangle}.$$

Example 2.1. We observe that 1-tensors can be identified with vectors $A \in \mathbb{R}^m$ by defining $A(x) = \langle A, x \rangle$. (Strictly speaking the vector A is the *holor* of the tensor defined this way.) The inner product is the usual inner product in \mathbb{R}^m , and the Frobenius norm $\|A\|_F = \|A\|_2$. Likewise 2-tensors can be identified with matrices $A \in \mathbb{R}^{m \times m}$ by defining $A(x, y) = \langle Ax, y \rangle$. Symmetric matrices $A = A^T$ give rise to symmetric 2-tensors. The inner product is $\langle A, B \rangle = \sum_{i,j} A_{ij} B_{ij}$ and $\|A\|_F$ is the matrix Frobenius norm.

Symmetric tensor fields Let $u : \Omega \rightarrow \text{Sym}^k(\mathbb{R}^m)$ for a domain $\Omega \subset \mathbb{R}^m$. We then set

$$\|u\|_{F,p} := \left(\int_{\Omega} \|u(x)\|_F^p dx \right)^{1/p} \quad (p \in [1, \infty)), \quad \text{and} \quad \|u\|_{F,\infty} := \text{ess sup}_{x \in \Omega} \|u(x)\|_F$$

as well as the abbreviation $\|\cdot\| = \|\cdot\|_{F,2}$, and define the spaces

$$L^p(\Omega; \text{Sym}^k(\mathbb{R}^m)) = \{u : \Omega \rightarrow \text{Sym}^k(\mathbb{R}^m) \mid \|u\|_{F,p} < \infty\}, \quad (p \in [1, \infty]).$$

For $u \in C^1(\Omega; \text{Sym}^k(\mathbb{R}^m))$, $k \geq 1$, we define the divergence $\text{div } u \in C(\Omega; \text{Sym}^{k-1}(\mathbb{R}^m))$ by contraction as

$$\begin{aligned} [\text{div } u(x)](e_{i_2}, \dots, e_{i_k}) &:= \sum_{i_1=1}^m \partial_{i_1} [x \mapsto u(x)(e_{i_1}, \dots, e_{i_k})] \\ &= \sum_{i_1=1}^m \langle e_{i_1}, \nabla u(\cdot)(e_{i_1}, \dots, e_{i_k}) \rangle. \end{aligned} \tag{2.1}$$

Observe the operator div preserves symmetricity.

Example 2.2. If $u \in C^1(\Omega; \mathbb{R}^m) = C^1(\Omega; \mathcal{T}^1(\mathbb{R}^m))$ is a vector field, then $\text{div } u(x) = \sum_{i=1}^m \partial_i u_i(x)$ is the usual divergence. If $u \in C^1(\Omega; \mathcal{T}^2(\mathbb{R}^m))$, then $[\text{div } u(x)]_j = \sum_{i=1}^m \partial_i u_{ij}(x)$. That is, we take columnwise the divergence of a vector field.

The symmetrised gradient Denoting by X^* the continuous linear functionals on the topological vector space X , we now define the symmetrised distributional gradient

$$Eu \in [C_c^\infty(\Omega; \text{Sym}^{k+1}(\mathbb{R}^m))]^*$$

of $u \in L^1(\Omega; \text{Sym}^k(\mathbb{R}^m))$ by

$$Eu(\varphi) := - \int_{\Omega} \langle u(x), \text{div } \varphi(x) \rangle dx, \quad (\varphi \in C_c^\infty(\Omega; \text{Sym}^{k+1}(\mathbb{R}^m))).$$

Let us also define the ‘‘symmetric Frobenius unit ball’’

$$V_{F,s}^k := \{\varphi \in C_c^\infty(\Omega; \text{Sym}^k(\mathbb{R}^m)) \mid \|\varphi\|_{F,\infty} \leq 1\}.$$

If $\sup\{Eu(\varphi) \mid \varphi \in V_{F,s}^{k+1}\} < \infty$, then Eu is a measure [10, §4.1.5]. Indeed, for our purposes it suffices to define a tensor-valued measure $\mu \in \mathcal{M}(\Omega; \text{Sym}^k(\mathbb{R}^m))$ as a linear functional $\mu \in [C_c^\infty(\Omega; \text{Sym}^k(\mathbb{R}^m))]^*$ bounded in the sense that the total variation norm

$$\|\mu\|_{F,\mathcal{M}(\Omega; \text{Sym}^k(\mathbb{R}^m))} := \sup\{\mu(\varphi) \mid \varphi \in V_{F,s}^k\} < \infty.$$

Example 2.3. For smooth vector fields $u \in C^\infty(\Omega; \mathbb{R}^m)$, we have $Eu = \mathcal{E}u\mathcal{L}^m$, where $\mathcal{E}u(x) := (\nabla u(x) + (\nabla u(x))^T)/2$ is the pointwise symmetrisation of the usual gradient, and \mathcal{L}^m denotes the Lebesgue measure; cf. [20]. For scalar fields $u \in L^1(\Omega)$, the symmetrised gradient is the usual distributional gradient, $Eu = Du$.

Miscellaneous notation We denote by

$$\delta_M(x) := \begin{cases} 0, & x \in M, \\ \infty, & x \notin M, \end{cases}$$

the indicator function of a set M in the sense of convex analysis, and particularly by

$$\delta_{\geq 0}(u) := \begin{cases} 0, & u(x) \text{ is positive semi-definite for a.e. } x \in \Omega, \\ \infty, & \text{otherwise,} \end{cases}$$

the indicator function of the pointwise positive semi-definite cone. We use the notation $u \geq 0$ for pointwise a.e. positive semi-definite $u \in L^1(\Omega; \text{Sym}^2(\mathbb{R}^m))$.

3 Total generalised variation of tensor fields

We now develop first- and second-order total generalised variation for tensor fields.

Total deformation (TGV¹) We define the first-order total generalised variation or *total deformation* of a tensor field $u \in L^1(\Omega; \text{Sym}^k(\mathbb{R}^m))$ simply by

$$\text{TD}(u) := \|Eu\|_{F, \mathcal{M}(\Omega; \text{Sym}^{k+1}(\mathbb{R}^m))} = \sup_{\varphi \in V_{F,s}^{k+1}} \int_{\Omega} u(x) \operatorname{div} \varphi(x) \, dx.$$

Observe that we bound φ pointwise by the Frobenius norm. The reason for this is that we desire rotation-invariance: for details see [23].

As already remarked, in the case of scalar fields $u \in L^1(\Omega; \text{Sym}^0(\mathbb{R}^m)) \sim L^1(\Omega)$, i.e., $k = 0$, we have $Eu = Du$, so that $\text{TD} = \text{TV}$. For $k = 1$, i.e., vector fields $u \in L^1(\Omega; \text{Sym}^1(\mathbb{R}^m)) \sim L^1(\Omega; \mathbb{R}^m)$, the functional TD agrees with the total deformation of [20]. We can also define tensorial total variation per $\text{TV}(u) := \|Du\|_{F, \mathcal{M}(\Omega; \mathcal{T}^{k+1}(\mathbb{R}^m))}$ where Du is the usual distributional derivative of u . This is done in [23], where we observe that TD provides better regularisation results.

We also denote $\text{TGV}_{\alpha}^1 := \alpha \text{TD}$, because total deformation is the first-order version of TGV. Moreover, we consider a fixed linear mapping $A : \text{Sym}^2(\mathbb{R}^m) \rightarrow X$ for a Hilbert space X and denote, again, by \mathcal{A} the corresponding pointwise superposition operator. The problem (P) may thus be written in inf-sup form

$$\min_v \sup_{\xi} G(v) + \langle v, \mathcal{K}^* \xi \rangle - F^*(\xi) \quad (\text{S})$$

with the unknowns

$$\begin{aligned} v &:= u \in L^1(\Omega; \text{Sym}^2(\mathbb{R}^m)), \\ \xi &:= \varphi \in C_c^{\infty}(\Omega; \text{Sym}^3(\mathbb{R}^m)), \end{aligned}$$

and the functionals

$$\begin{aligned} G(u) &:= \frac{1}{2} \|f - \mathcal{A}u\|^2 + \delta_{\geq 0}(u), \\ \mathcal{K}^* \varphi &:= -\operatorname{div} \varphi, \quad \text{and} \\ F^*(\varphi) &:= \delta_{\alpha V_{F,s}^3}(\varphi). \end{aligned} \quad (\text{S-TD})$$

The conjugate-like notation \mathcal{K}^* will be justified in Section 4.

Alternatively, we may introduce, as for instance in [5], an additional dual variable λ , and write

$$\frac{1}{2} \|f - \mathcal{A}u\|^2 = \sup_{\lambda} \langle \lambda, \mathcal{A}u - f \rangle^2 - \frac{1}{2} \|\lambda\|^2. \quad (3.1)$$

Then with the unknowns

$$\begin{aligned} v &:= u \in L^1(\Omega; \text{Sym}^2(\mathbb{R}^m)), \\ \xi &:= (\varphi, \lambda) \in C_c^{\infty}(\Omega; \text{Sym}^3(\mathbb{R}^m)) \times L^1(\Omega; \text{Sym}^2(\mathbb{R}^m)) \end{aligned}$$

we set

$$\begin{aligned} G(u) &:= \delta_{\geq 0}(u), \\ \mathcal{K}^*(\varphi, \lambda) &:= (-\operatorname{div} \varphi + \mathcal{A}^* \lambda), \quad \text{and} \\ F^*(\varphi) &:= \delta_{\beta V_{F,s}^3}(\varphi) + \frac{1}{2} \|\lambda\|^2 + \langle f, \lambda \rangle. \end{aligned} \quad (\text{S-TD-alt})$$

Second-order total generalised variation (TGV²) Second-order total generalised variation, introduced in [6], tends to avoid [16] the stair-casing effect of first-order regularisation models. Given parameters $\vec{\alpha} = (\alpha, \beta) > 0$, we now define it for tensor fields $u \in L^1(\Omega; \text{Sym}^k(\mathbb{R}^m))$ by the differentiation cascade

$$\text{TGV}_{\vec{\alpha}}^2(u) := \min_{w \in L^1(\Omega; \text{Sym}^{k+1}(\mathbb{R}^m))} \alpha \|Eu - w\|_{F, \mathcal{M}(\Omega; \text{Sym}^{k+1}(\mathbb{R}^m))} + \beta \|Ew\|_{F, \mathcal{M}(\Omega; \text{Sym}^{k+2}(\mathbb{R}^m))}. \quad (3.2)$$

It is shown in [7, 8] that this formulation is for scalar fields $k = 0$ equivalent to the original dual-ball formulation in [6]. We choose to use the formulation (3.2) directly because it is more practical for numerical realisation, simplifying constraints through a kind of variable splitting.

The problem (P) for $\ell = 2$ may again be written in the inf-sup form (S) with the unknowns

$$\begin{aligned} v &:= (u, w) \in L^1(\Omega; \text{Sym}^2(\mathbb{R}^m)) \times L^1(\Omega; \text{Sym}^3(\mathbb{R}^m)), \quad \text{and} \\ \xi &:= (\varphi, \psi) \in C_c^\infty(\Omega; \text{Sym}^3(\mathbb{R}^m)) \times C_c^\infty(\Omega; \text{Sym}^4(\mathbb{R}^m)), \end{aligned}$$

and the functionals

$$\begin{aligned} G(u, w) &:= \frac{1}{2} \|f - \mathcal{A}u\|^2 + \delta_{\geq 0}(u), \\ \mathcal{K}^*(\varphi, \psi) &:= (-\text{div } \varphi, -\varphi - \text{div } \psi), \quad \text{and} \\ F^*(\varphi, \psi) &:= \delta_{\alpha V_{F,s}^3}(\varphi) + \delta_{\beta V_{F,s}^4}(\psi). \end{aligned} \tag{S-TGV²}$$

Alternatively, we may do as in (3.1), and with the unknowns

$$\begin{aligned} v &:= (u, w) \in L^1(\Omega; \text{Sym}^2(\mathbb{R}^m)) \times L^1(\Omega; \text{Sym}^3(\mathbb{R}^m)), \quad \text{and} \\ \xi &:= (\varphi, \psi, \lambda) \in C_c^\infty(\Omega; \text{Sym}^3(\mathbb{R}^m)) \times C_c^\infty(\Omega; \text{Sym}^4(\mathbb{R}^m)) \times L^1(\Omega; \text{Sym}^2(\mathbb{R}^m)) \end{aligned}$$

write

$$\begin{aligned} G(u, w) &:= \delta_{\geq 0}(u), \\ \mathcal{K}^*(\varphi, \psi, \lambda) &:= (-\text{div } \varphi + \mathcal{A}^* \lambda, -\varphi - \text{div } \psi), \quad \text{and} \\ F^*(\varphi, \psi, \lambda) &:= \delta_{\alpha V_{F,s}^3}(\varphi) + \delta_{\beta V_{F,s}^4}(\psi) + \frac{1}{2} \|\lambda\|^2 + \langle f, \lambda \rangle. \end{aligned} \tag{S-TGV²-alt}$$

4 The numerical method

We now move on to discuss the algorithmic aspects of the solution of the regularisation problems above. We do this through the min-sup formulations.

Discretisation and the algorithm We intend to apply the Chambolle-Pock algorithm [9] to the inf-sup problem (S) for the functionals (S-TD), (S-TD-alt), (S-TGV²), and (S-TGV²-alt). This can be done after we discretise the original problem first; for the infinite-dimensional problem the (pre)conjugate \mathcal{K} of \mathcal{K}^* is not well-defined. We represent each tensor field f , u , w , φ and ψ with values on a uniform rectangular grid Ω_h of cell width $h > 0$, and discretise the operator div by forward differences with zero boundary conditions as div_h . We choose not to use central differences, because it tends to cause oscillation in this problem. The operator \mathcal{A}^* is discretised in the straightforward manner as pointwise application of A^* on the grid points. This yields the discretised version \mathcal{K}_h^* of the operator \mathcal{K}^* . We then take $\mathcal{K}_h := (\mathcal{K}_h^*)^*$ as the discrete conjugate of \mathcal{K}_h^* .

We also like to employ the acceleration strategies for the primal-dual algorithm in [9]. Assuming that the operator A has full rank, observe that G for TD as given in (S-TD), is uniformly convex in u in the sense that there exists a factor $\gamma > 0$ (dependent on A) such that for any u' it holds

$$G(u') - G(u) \geq \langle z, u' - u \rangle + \frac{\gamma}{2} \|u - u'\|^2 \quad \text{for all } z \in \partial G(u), x' \in X.$$

This implies that acceleration can be employed. The alternative formulations (S-TD-alt) for TD, and (S-TGV²-alt) for TGV², however, are not uniformly convex, so $\gamma = 0$. For TGV², the functional G as given in (S-TGV²), is uniformly convex in u , but not in w . We should therefore take $\gamma = 0$ in the algorithm below, although in numerical practise taking γ as the factor of uniform convexity with respect to u alone, sometimes offers faster convergence, at other times however failing to exhibit convergence.

Algorithm 4.1. Perform the steps:

1. Pick $\tau_0, \sigma_0 > 0$ satisfying $\tau_0 \sigma_0 \|\mathcal{K}_h\|^2 < 1$, as well as initial iterates (v^0, ξ^0) . Set $\bar{v}^0 = v^0$. Let γ be the factor of uniform convexity of G .
2. For $i = 0, 1, 2, \dots$, repeat until a stopping criterion is satisfied.

$$\begin{aligned}\xi^{i+1} &:= (I + \sigma_i \partial F^*)^{-1}(\xi^i + \sigma_i \mathcal{K}_h \bar{v}^i) \\ v^{i+1} &:= (I + \tau_i \partial G)^{-1}(v^i - \tau_i \mathcal{K}_h^* \xi^{i+1}) \\ \theta_i &:= (1 + 2\gamma\tau_i)^{-1/2}, \tau_{i+1} := \theta_i \tau_i, \sigma_{i+1} := \sigma_i / \theta_i \\ \bar{v}^{i+1} &:= v^{i+1} + \theta_i (v^{i+1} - v^i).\end{aligned}$$

Here, $(I + \sigma \partial F^*)^{-1}$ and $(I + \tau \partial G)^{-1}$ are resolvent mappings associated with the subgradients of F^* and G , respectively, for details see [17], for instance.

Calculating the resolvents We now study how to calculate for (S-TGV²) the resolvents $(I + \tau \partial G)^{-1}(v)$ and $(I + \sigma \partial F^*)^{-1}(\xi)$, which are needed to obtain ξ^{i+1} and v^{i+1} in Algorithm 4.1. The solution for (S-TD) is then a simplified version. For (S-TD-alt) and (S-TGV²-alt) the resolvent $(I + \tau \partial G)^{-1}(v)$ is likewise a simplified version of the case considered below, while $(I + \sigma \partial F^*)^{-1}(\xi)$ amounts in each case to simple pointwise operations on ξ , which have already been discussed in the literature; for instance, see [23, 5].

The resolvents have the expressions

$$(I + \tau \partial G)^{-1}(v) = \arg \min_y \left\{ \frac{\|v - y\|^2}{2\tau} + G(y) \right\},$$

where for pairs $v = (u, w)$ we have to take $\|v\|^2 = \|u\|^2 + \|w\|^2$. The efficient realisation of Algorithm 4.1 depends on the efficient realisation of these minimisation problems. As discussed in [23], calculation of $(I + \sigma \partial F^*)^{-1}(\varphi, \psi)$ reduces to pointwise projection of φ to the ball $\alpha V_{F,s}^3$ and of ψ to $\alpha V_{F,s}^4$. Moreover, for (S-TD-alt) and (S-TGV²-alt), λ is mapped to $(\lambda - \sigma f)/(1 + \sigma)$. The calculation of $(I + \tau \partial G)^{-1}(u, w)$ for (S-TD) and (S-TGV²) is more demanding. Recalling that

$$G(u, w) = \frac{1}{2} \|f - Au\|_2^2 + \delta_{\geq 0}(u),$$

we find that

$$(u, w) = (I + \tau \partial G)^{-1}(v, q)$$

satisfy $w = q$, and that $u(x)$ for each $x \in \Omega$ is the solution of

$$(I + \tau A^* A)u(x) + N_{\geq 0}(u(x)) \ni v(x) + \tau A^* f(x). \quad (4.1)$$

Here $N_{\geq 0}(u(x))$ is the normal cone of the positive semi-definite cone at $u(x)$. If $A = I$, this simplifies to

$$(1 + \tau)u(x) + N_{\geq 0}(u(x)) \ni v(x) + \tau f(x),$$

so that

$$u(x) = P_{\geq 0} \left(\frac{v(x) + f(x)\tau}{1 + \tau} \right).$$

The projection $P_{\geq 0}$ to the positive definite cone can be calculated by performing eigen-decomposition using QR algorithm, see [12] for details. This is also true for (S-TD-alt) and (S-TGV²-alt) in which case the resolvent is just the projection on the positive definite cone. If $A \neq I$, then writing $M = I + \tau A^* A$ and $C = v(x) + \tau A^* f(x)$, the problem (4.1) can be stated in terms of $X = u(x) \in \text{Sym}^2(\mathbb{R}^m)$ and a dual variable $S \in \text{Sym}^2(\mathbb{R}^m)$ as

$$MX - S = C, (XS + SX)/2 = 0, X, S \geq 0. \quad (4.2)$$

This system of equations can be approximately solved in a low number of steps with the novel primal-dual interior point algorithm of [22].

Finally, we remark that when we do not include the positivity constraint (P), then the projections above may be skipped, and instead of (4.1), we simply solve $u(x)$ from

$$Mu(x) = C.$$

Stopping criterion For (S-TD) and (S-TD-alt), we use the fractional duality gap as stopping criterion. Namely we stop execution of Algorithm 4.1 if a preset maximum number of iterations is reached, or for a given target fractional duality gap $\rho \in (0, 1)$, we have $D^i/D^0 < \rho$, where the duality gap

$$D^i := F(\mathcal{K}_h v^i) + G(v^i) + G^*(-\mathcal{K}_h^* \xi^i) + F^*(\xi^i). \quad (4.3)$$

For (S-TGV²) and (S-TGV²-alt) we cannot use this as such, because often simply $D^i = \infty$ due to the independence of G from w . In [23] we have shown that by including in G the term $M\|\hat{w}\|_{F,1}$, for an *a posteriori* chosen $M > 0$, we can use the resulting “pseudo-duality gap” as a stopping criterion.

5 Numerical results

Experimental setup We apply the regularisation models discussed above to a clinical in-vivo diffusion tensor image of a human brain. The measurements for our test data set were performed on a clinical 3T system (Siemens Magnetom TIM Trio, Erlangen, Germany), using a 32 channel head coil. Written informed consent was obtained from all volunteers before the examination. A 2D diffusion weighted single shot EPI sequence with diffusion sensitising gradients applied in 12 independent directions ($b = 1000\text{s/mm}^2$) and an additional reference scan without diffusion was used with the following sequence parameters: TR = 7900ms, TE = 94ms, flip angle 90° , matrix size 128×128 , 60 slices with a slice thickness of 2mm, in plane resolution $1.95\text{mm} \times 1.95\text{mm}$, 4 averages, GRAPPA acceleration factor 2. The exact measurement time of the entire sequence is 7 minutes and 22 seconds. Prior to the reconstruction of the diffusion tensor, eddy current correction was performed with FSL [19].

We perform computations only on slice 23 of the data set. As a ground truth, we construct the diffusion tensor image g_0 by the Stejskal-Tanner equation (1.1) from all the 52 (4 averages of 12+1) DWI measurements $\hat{s} = (s_0, \dots, s_{51})$. We then take only the 7 measurements $s = (s_0, s_2, s_4, s_5, s_6, s_8, s_{11})$, containing one zero diffusion gradient and six independent gradients. The six independent gradients are chosen among all the 12 to minimise the condition number of the operator A^*A , where A is constructed by (1.3). From the *reduced data set* s we construct a lower-quality image g by the Stejskal-Tanner equation. This kind of reduction of the number of measurements would bring the scanning time from the 7 minutes and 22 seconds to below 2 minutes, which would be beneficial to reduce artefacts from patient motion and improve patient comfort.

We denoise g directly by minimising $\|g - u\|^2$ with both TGV² and TD regulariser, with and without the positivity constraint. We call this the “direct denoising problem”, as we are directly minimising the distance between two tensor fields subject to regularisation and constraints. We also denoise g “indirectly” by minimising $\|f - Au\|^2$, where f and A are as in (1.2) and (1.3), respectively. We again use both TGV² and TD as regulariser, both with and without the positivity constraint. We call this the “raw DWI-based denoising problem”, because the fidelity function incorporates the DWI measurements through the Stejskal-Tanner equation.

We compute the denoising result of every model for all the regularisation parameters

$$\alpha \in \{0.0003, 0.0006, 0.0009, 0.0012, 0.0015, 0.0018, 0.0021, 0.0024, 0.0027, \\ 0.00003, 0.00006, 0.00009, 0.00012, 0.00015, 0.00018, 0.00021, 0.00024, 0.00027\}.$$

For TGV² we take $\beta = \alpha$. This range was chosen by trial and error to yield qualitatively good results and also be large enough such that when we calculate for each α the Frobenius norm error $\|\chi_{\text{brain}}(u - g_0)\|_{F,2}$ to the ground-truth g_0 , then the optimal α , yielding smallest error, lies inside this

range, not on the boundary. Here $\chi_{\text{brain}} : \Omega \rightarrow \{0, 1\}$ is the brain mask, defined as excluding points where the average DWI signal intensity is less than 10% of the mean over the whole image. In the comparisons that we now get into, for each model we report the result for the optimal choice of α . All the reported error scores also use the mask χ_{brain} , and the corresponding area has been masked out in the visualisations, to avoid displaying distractive noise outside the imaged brain.

Performance considerations between formulations We use for all Algorithm 4.1 for all computations. For the situation $A = I$, i.e., the direct denoising problem, the formulations (S-TD) and (S-TGV²) are faster convergent than the formulations (S-TD-alt) and (S-TGV²-alt) with the extra dual variable λ . Indeed, by the uniform convexity of (S-TD), Algorithm 4.1 has $O(1/N^2)$ convergence rate [9]. This still holds for $A \neq I$, i.e., the raw DWI-based denoising problem, but in practise we have observed that if the operator A is far from orthogonal, although initial convergence is fast, the algorithm quickly becomes slowly convergent without having reached a high-quality solution. This happens both with and without the positivity constraint. In most situations we have found the target fractional duality gap $\rho = 0.001$ to offer reasonable solutions. This is also the case for the current “optimal” choice of the six independent diffusion gradients (with condition number ~ 6.8 for A^*A). If, however, we picked the 6 first independent diffusion gradients, i.e., set $s = (s_0, \dots, s_7)$ (which yields condition number ~ 33.6 for A^*A), we have found that we would need at least $\rho = 0.00001$ to get reasonable solutions with the formulations (S-TD) and (S-TGV²). This results in higher computational times. We note, however, that although the effect is less extreme, the formulations (S-TD-alt) and (S-TGV²-alt) also exhibit diminished convergence speeds in this case ($\rho = 0.001$ giving red tone in the principal eigenvector plots), suggesting that the whole raw DWI-based denoising problem has wider problems with A^*A not being the identity.

In the positivity constrained case there are also some additional numerical difficulties with the formulations (S-TD) and (S-TGV²). Although the interior point algorithm of [22] can in few iterations find an approximate solution (fractional duality gap < 0.001) to (4.2), and this is sufficient to obtain good solutions with Algorithm 4.1, such approximate solutions are not sufficiently good for calculating the fractional duality gap (4.3) for use as a stopping criterion in Algorithm 4.1. Namely, in order to calculate the duality gap, we need to compute a high-quality approximation of the conjugate

$$G^*(q) := \sup_u \{\langle q, u \rangle - G(u)\}.$$

This also involves the solution of (4.2). When we calculate the solution only approximately, G^* and then consequently the duality gap ends up computed too small, causing early termination of the Algorithm 4.1, or none at all, when the initial duality gap is computed too small. Thus we presently have to compute the solution to (4.2) very exactly, which results in rather long running times. The higher condition number of A^*A also increases the running times, as the complexity of the interior point method depends on the condition number. Alternatively we should not use the duality gap as stopping criterion, and simply run Algorithm 4.1 for a fixed number of iterations.

The above performance considerations have led us to the following conclusions. For the direct denoising problem, we employ the formulations (S-TD) and (S-TGV²). For the raw DWI-based denoising problem, we employ the alternative problem formulations (S-TD-alt) and (S-TGV²-alt) if A^*A is far from the identity. For A from the present experimental setup (with the optimal choice of the independent diffusion gradients), the Chambolle-Pock algorithm however converges fast for the formulations (S-TD) and (S-TGV²), so we use them. Due to the problems with the exact calculation of the duality gap, the run time of the method (in seconds) is however not comparable for the constrained raw DWI-based fidelity to the other fidelities.

Quantitative results We set the target fractional duality gap $\rho = 0.001$ and a maximum iteration count 5000. With this choice, error scores for the computational results for the optimal α are reported in Table 1. In addition to the Frobenius norm error $\|\chi_{\text{brain}}(g_0 - u)\|_{F,2}$, we report the error

Table 1: Computational results for all of the evaluated models. The parameter α that provides smallest Frobenius norm error has been selected for each model. This error is reported (middle column) along with the error of the fractional anisotropies and iteration count. The errors reported exclude points outside the brain mask.

Model	Optimal α	Frobenius error	FA error	Iterations
noisy data g		0.03195	9.96522	
direct unconstr. TD	0.00030	0.02563	7.43823	71
direct unconstr. TGV ²	0.00030	0.02480	<u>7.30000</u>	69
raw unconstr. TD	0.00024	0.02535	8.57360	23
raw unconstr. TGV ²	0.00024	0.02554	8.00827	58
direct pos.def. TD	0.00024	0.02281	7.88564	74
direct pos.def. TGV ²	0.00024	0.02183	7.82056	63
raw pos.def. TD	0.00018	<u>0.02145</u>	7.84746	22
raw pos.def. TGV ²	0.00018	0.02159	7.49770	51

$\|\chi_{\text{brain}}(\text{FA}_{g_0} - \text{FA}_u)\|_{L^2(\Omega)}$ in the fractional anisotropies

$$\text{FA}_u(x) = \left(\sum_{i=1}^m (\lambda_i - \bar{\lambda})^2 \right)^{1/2} \left(\sum_{i=1}^m \lambda_i^2 \right)^{-1/2} \in [0, 1], \quad (x \in \Omega).$$

Here $\lambda_1, \dots, \lambda_m$ denote the eigenvalues of $u(x)$, and $\bar{\lambda} = \sum_{i=1}^m \lambda_i / m$ their average.

As we see, raw DWI-based positivity constrained TGV² and TD perform the best with regard to the Frobenius norm. Direct unconstrained TGV² has the best error score for fractional anisotropy norm (which was not used as criterion for the choice of α).

The differences between TD and TGV² are in each case minor. In the following paragraphs we will therefore only study with regard to TGV² the raw DWI-based versus direct tensor field denoising, and positivity constrained versus unconstrained denoising. Later we will see that increasing the noise level introduces significant differences between TD and TGV².

Raw versus direct, constrained versus unconstrained In Figure 1 we display the colour-coded results for TGV² with and without the positivity constraint for both the raw and direct fidelity functions. Figure 2 includes the corresponding colour-coded error plot, and Figure 3 displays the *corpus callosum* region in detail, with greyscale coding of the fractional anisotropy, superimposed by the principal eigenvectors. The location of this region in the brain is indicated in Figure 1. Finally, Figure 4 contains a visualisation of the positivity of the tensors.

As we see, direct unconstrained denoising has clearly the worst results, greatly over-smoothed. We note from Figure 4, however, that most negative eigenvalues within the brain mask are removed even by unconstrained denoising. The effect of the constraint therefore lies primarily elsewhere.

We see, moreover, that the constrained direct denoising result is very noisy in the central region of the brain, while the raw denoising better restores the original data. The error plot of Figure 2 indicates that direct denoising leaves more errors in fractional anisotropy in this area. Studying the *corpus callosum* in Figure 3, we see that direct denoising however better restores the dark (low-anisotropy) region under the high-anisotropy (near-white) corpus callosum. Raw DWI-based denoising however has better performance with respect to the directions of the principal eigenvectors.

TGV² versus TD, adding more noise The differences in the computational results in Table 1 between TD and TGV² are insignificant, and would be unobservable by the eye in the visualisations. We have therefore performed additional computations with a higher noise level. We still take $s = (s_0, s_2, s_4, s_5, s_6, s_8, s_{11})$, and apply additional Rician noise of parameter $\sigma = 10$ to these signals. We then solve the tensor field g from (1.1).

Table 2: Computational results for all of the evaluated models in the case with additional Rician noise of parameter $\sigma = 10$. The parameter α that provides smallest Frobenius norm error has been selected for each model. This error is reported (middle column) along with the error of the fractional anisotropies and iteration count. The errors reported exclude points outside the brain mask.

Model	Optimal α	Frobenius error	FA error	Iterations
noisy data g		0.05067	17.33472	
direct unconstr. TD	0.00060	0.03383	10.12227	56
direct unconstr. TGV ²	0.00060	0.03261	9.57858	123
raw unconstr. TD	0.00060	0.03007	11.13348	21
raw unconstr. TGV ²	0.00060	0.03018	9.68956	108
direct pos.def. TD	0.00060	0.03168	9.99513	58
direct pos.def. TGV ²	0.00060	0.03032	9.43669	113
raw pos.def. TD	0.00030	<u>0.02628</u>	9.50354	20
raw pos.def. TGV ²	0.00030	0.02681	<u>9.12207</u>	69

As in the earlier lower-noise test case, we apply all the eight different denoising models. The error scores are reported in Table 2. In Figure 5 we moreover display the colour-coded results for TGV² and TD with the positivity constraint for both the raw and direct fidelity functions. Figure 6 includes the corresponding colour-coded error plot, and Figure 7 displays the *corpus callosum* region in detail, with greyscale coding of the fractional anisotropy, superimposed by the principal eigenvectors. The location of this region in the brain is indicated in Figure 5.

Clearly raw DWI-based positivity constrained TGV² and TD performs the best with respect to the Frobenius norm, with other denoising models far behind. Raw DWI-based positivity constrained TGV² also performs significantly better than other methods with respect to the fractional anisotropy error. Studying Figure 5, we immediately notice that direct positivity-constrained TD exhibits some stair-casing. Study of the error plot in Figure 6, and the corpus callosum in Figure 7 indicates further higher errors of TD versus TGV².

6 Conclusions

We conclude from our studies that DWI-based denoising, incorporating the Stejskal-Tanner equation (1.1) in (P), offers improved results over the approach in [23], where we first solve the diffusion tensor field u from (1.1), and then denoise u . The effect of the positivity constraint is a more moot point. Per the error scores, in Table 1, it provides improved results, but visually this is hardly noticeable. Likewise is the situation with first-order versus second-order regularisation in the test case without additional noise. With additional noise applied to the measurements, we however start to experience stair-casing with first-order regularisation.

Acknowledgements

This research has been financially supported by the SFB research program F32 “Mathematical Optimization and Applications in Biomedical Sciences” of the Austrian Science Fund (FWF). The authors would moreover like to thank Karl Koschutnig (Department of Psychology, University of Graz) for providing the in-vivo brain DWI data.

References

1. M. Aksoy, C. Forman, M. Straka, S. Skare, S. Holdsworth, J. Hornegger and R. Bammer, *Real-time optical motion correction for diffusion tensor imaging.*, Magnetic Resonance in Medicine **66** (2011), 366–378,

doi:[10.1002/mrm.22787](https://doi.org/10.1002/mrm.22787).

2. V. Arsigny, P. Fillard, X. Pennec and N. Ayache, *Fast and simple computations on tensors with log-euclidean metrics.*, Technical Report 5584, INRIA (2005).
3. P. J. Basser and D. K. Jones, *Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review.*, NMR in Biomedicine **15** (2002), 456–467, doi:[10.1002/nbm.783](https://doi.org/10.1002/nbm.783).
4. R. L. Bishop and S. I. Goldberg, *Tensor Analysis on Manifolds*, Dover Publications, 1980, Dover edition.
5. K. Bredies, *Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty*, SFB-Report 2012-006, Karl-Franzens University of Graz (2012).
6. K. Bredies, K. Kunisch and T. Pock, *Total generalized variation*, SIAM J. Imaging Sci. **3** (2011), 492–526, doi:[10.1137/090769521](https://doi.org/10.1137/090769521).
7. K. Bredies, K. Kunisch and T. Valkonen, *Properties of L^1 -TGV²: The one-dimensional case*, J. Math. Anal Appl. (2012). Accepted.
8. K. Bredies and T. Valkonen, *Inverse problems with second-order total generalized variation constraints*, in: *Proceedings of SampTA 2011 – 9th International Conference on Sampling Theory and Applications*, Singapore, 2011.
9. A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision **40** (2011), 120–145, doi:[10.1007/s10851-010-0251-1](https://doi.org/10.1007/s10851-010-0251-1).
10. H. Federer, *Geometric Measure Theory*, Springer, 1969.
11. P. Fillard, X. Pennec, V. Arsigny and N. Ayache, *Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics*, IEEE Trans. Medical Imaging **26** (2007), 1472–1482.
12. G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 1996.
13. H. Gudbjartsson and S. Patz, *The Rician distribution of noisy MRI data*, Magnetic Resonance in Medicine **34** (1995), 910–914.
14. M. Herbst, J. Maclaren, M. Weigel, J. Korvink, J. Hennig and M. Zaitsev, *Prospective motion correction with continuous gradient updates in diffusion weighted imaging*, Magnetic Resonance in Medicine (2011), doi:[10.1002/mrm.23230](https://doi.org/10.1002/mrm.23230).
15. P. Kingsley, *Introduction to diffusion tensor imaging mathematics: Parts I-III*, Concepts in Magnetic Resonance Part A **28** (2006), 101–179, doi:[10.1002/cmr.a.20048](https://doi.org/10.1002/cmr.a.20048). [10.1002/cmr.a.20049](https://doi.org/10.1002/cmr.a.20049), [10.1002/cmr.a.20050](https://doi.org/10.1002/cmr.a.20050).
16. F. Knoll, K. Bredies, T. Pock and R. Stollberger, *Second order total generalized variation (TGV) for MRI.*, Magnetic Resonance in Medicine **65** (2011), 480–491, doi:[10.1002/mrm.22595](https://doi.org/10.1002/mrm.22595).
17. R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, 1998.
18. S. Setzer, G. Steidl, B. Popilka and B. Burgeth, *Variational methods for denoising matrix fields*, in: *Visualization and Processing of Tensor Fields*, Edited by J. Weickert and H. Hagen, Springer, 2009, 341–360.
19. S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. D. Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. D. Stefano, J. M. Brady and P. M. Matthews, *Advances in functional and structural MR image analysis and implementation as FSL.*, Neuroimage **23 Suppl 1** (2004), S208–S219, doi:[10.1016/j.neuroimage.2004.07.051](https://doi.org/10.1016/j.neuroimage.2004.07.051).
20. R. Temam, *Mathematical problems in plasticity*, Gauthier-Villars, 1985.
21. J.-D. Tournier, S. Mori and A. Leemans, *Diffusion tensor imaging and beyond*, Magnetic Resonance in Medicine **65** (2011), 1532–1556, doi:[10.1002/mrm.22924](https://doi.org/10.1002/mrm.22924).
22. T. Valkonen, *A method for weighted projections to the positive definite cone*, SFB-Report 2012-016, Karl-Franzens University of Graz (2012).
23. T. Valkonen and F. Knoll, *Total generalised variation in diffusion tensor imaging*, SFB-Report 2012-003, Karl-Franzens University of Graz (2012).

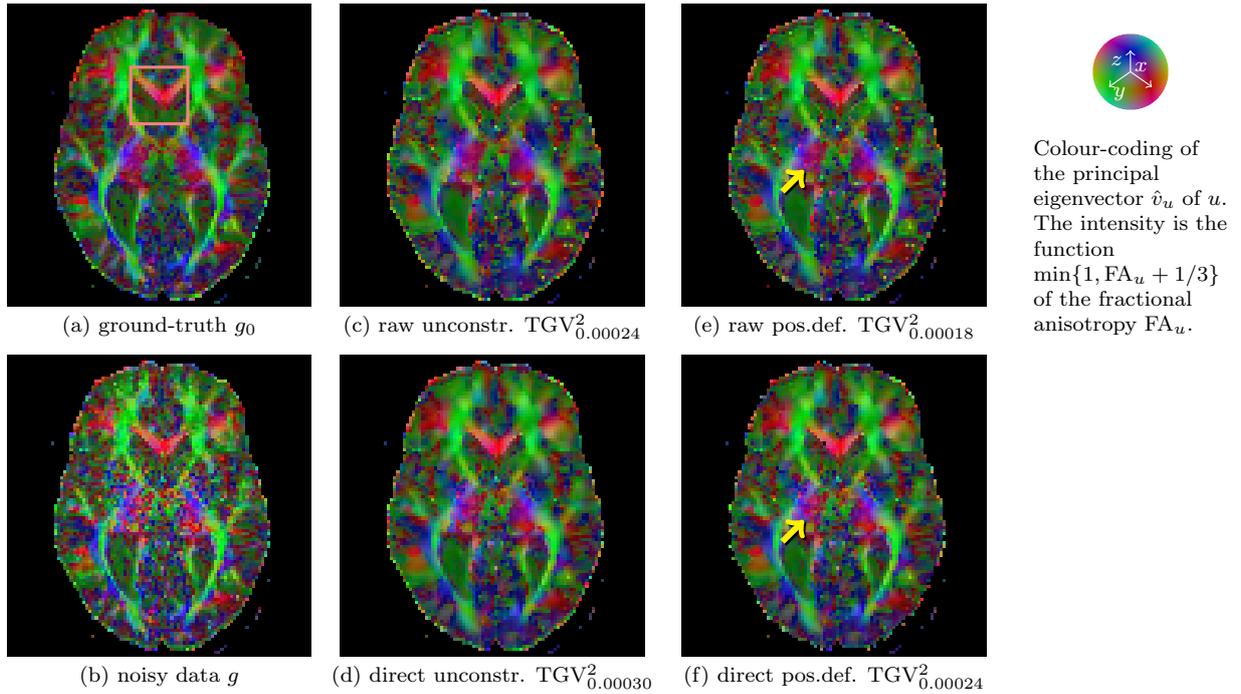


Figure 1: Colour-coded conventional reconstructions of the tensor field from full and reduced data, and results of TGV^2 denoising for raw and direct fidelity functions, constrained and unconstrained. The rectangle in (a) indicates the *corpus callosum* displayed in detail in Figure 3. Between (c) and (e) there is no significant visible difference. Direct denoising leaves “noise” in the centre, as indicated by the arrow in (f); compare (e).

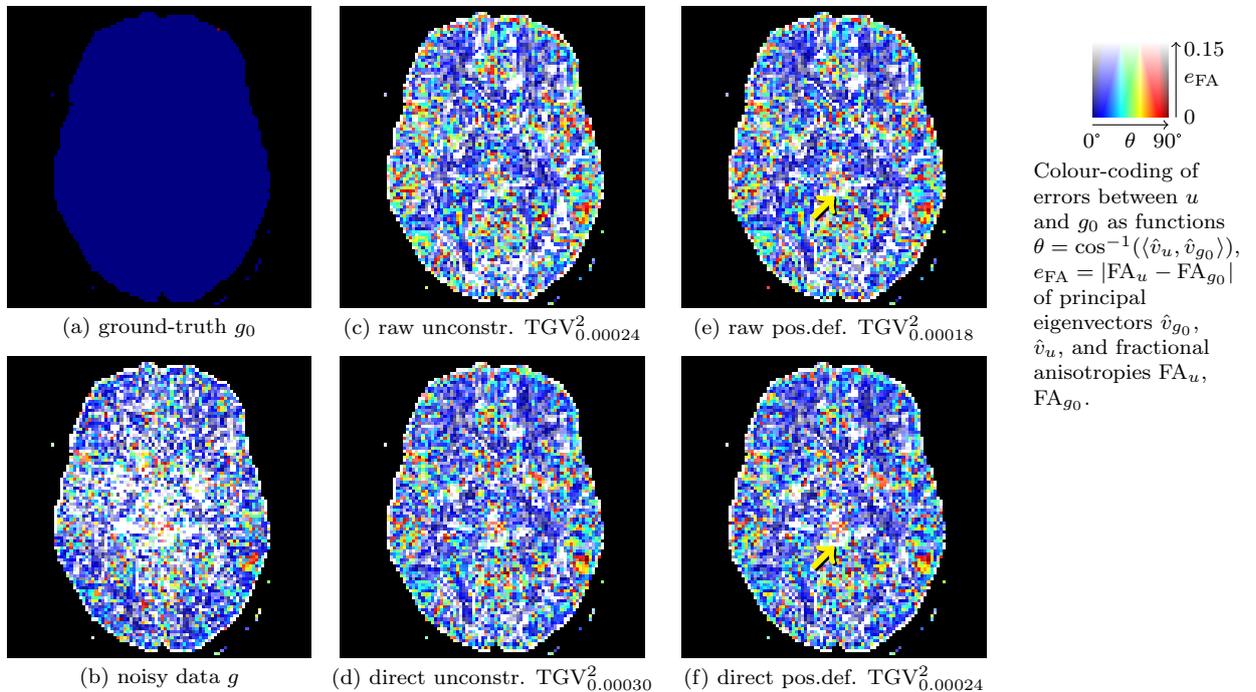


Figure 2: Colour-coded errors with respect to the ground-truth g_0 of TGV^2 denoising of reduced data for raw and direct fidelity functions, constrained and unconstrained. Errors in fractional anisotropy appear as white, and high errors in principal eigenvectors as red. Direct constrained denoising is worse at removing fractional anisotropy errors in the central region, as indicated by the arrows in (f); compare (e).

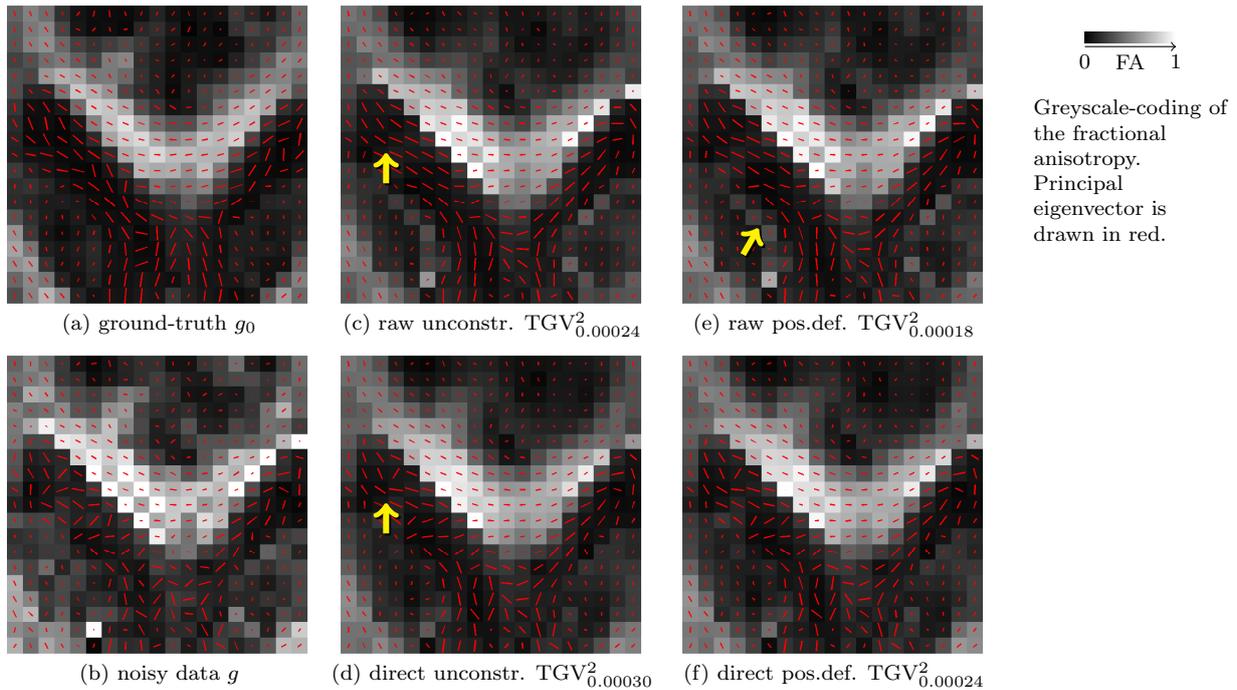


Figure 3: The *corpus callosum* region with fractional anisotropy in greyscale and principal eigenvector in red. Conventional reconstructions of the tensor field from full and reduced data and results of TGV^2 denoising of reduced data for raw and direct fidelity functions, constrained and unconstrained. In the low-anisotropy region under the corpus callosum, indicated by the arrows in (c) and (d), raw denoising restores the principal eigenvectors much more smoothly than direct denoising. It however introduces “patterns” in fractional anisotropy, as indicated by the arrow in (e).

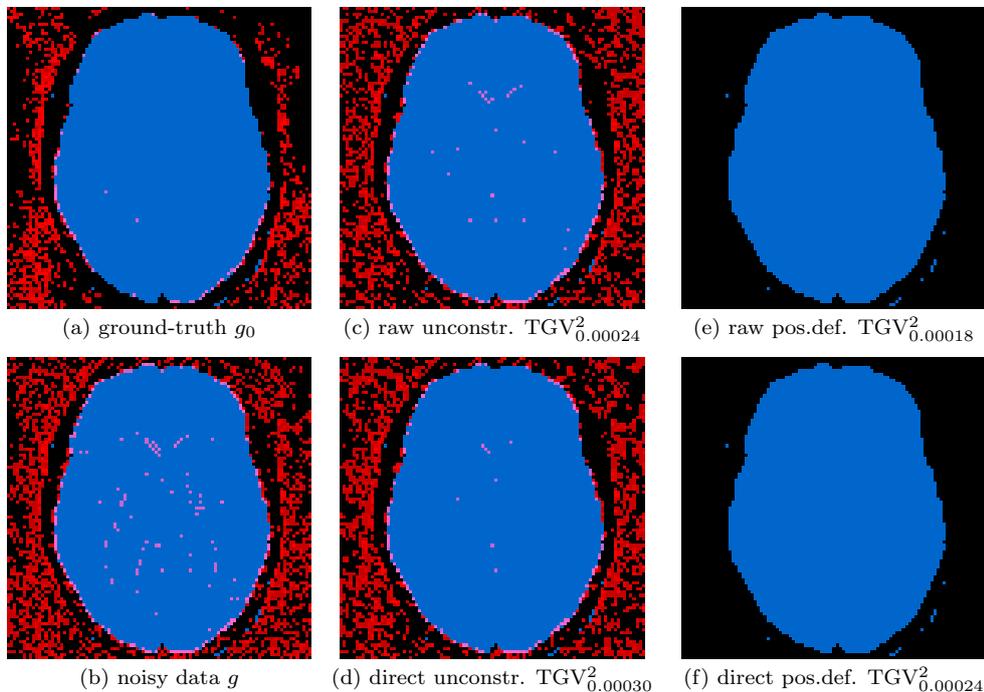


Figure 4: Visualisation of positivity. Conventional reconstructions of the tensor field from full and reduced data and computational results of TGV^2 denoising of reduced data for raw and direct fidelity functions, constrained and unconstrained. The light blue area indicates the brain mask, while red indicates negative eigenvalues. As we see, most negative eigenvalues within the brain mask in (c) (raw) and (d) (direct) are removed by TGV^2 denoising even without the positivity constraint, but not everything, as by constrained TGV^2 in (e) and (f).

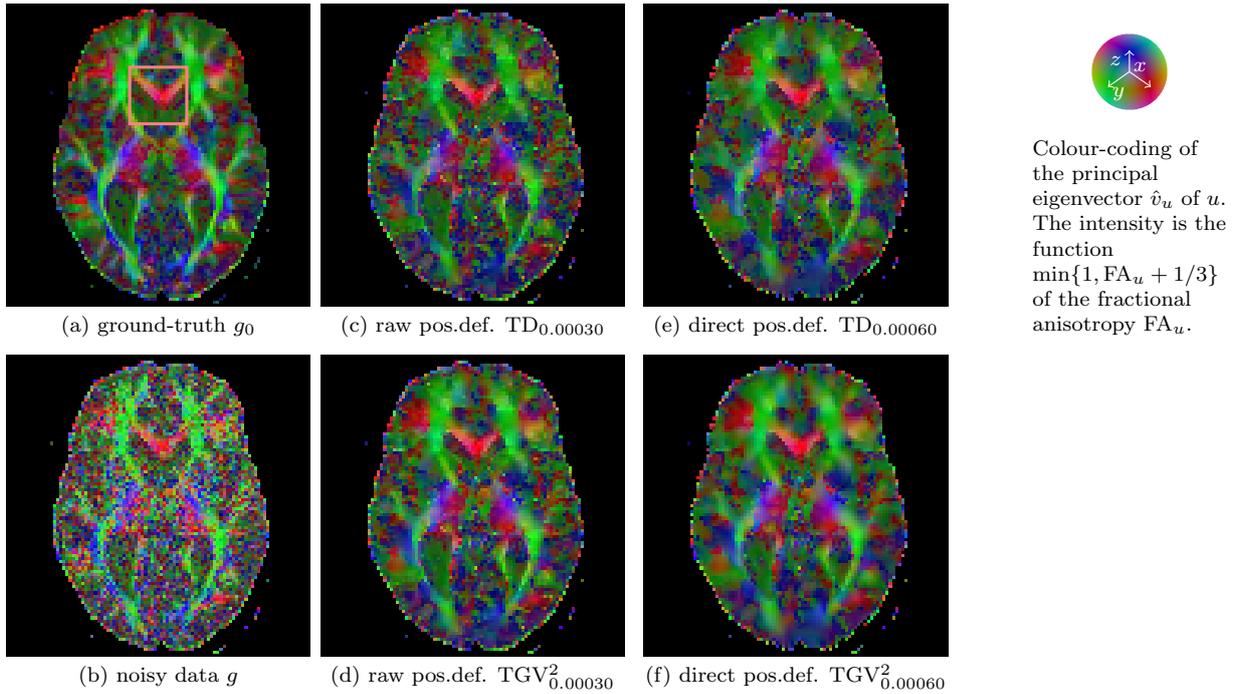


Figure 5: Colour-coded conventional reconstructions of the tensor field from full data and reduced data with additional Rician noise of parameter $\sigma = 10$ along with results of TGV^2 and TD denoising for raw and direct constrained fidelity functions. The rectangle in (a) indicates the *corpus callosum* displayed in detail in Figure 7. Observe significant stair-casing by TD in (e), while TGV^2 in (f) has become significantly smoothed. The differences between (c) and (d) (raw denoising) are lesser, but some stair-casing/roughness can be observed for TD.

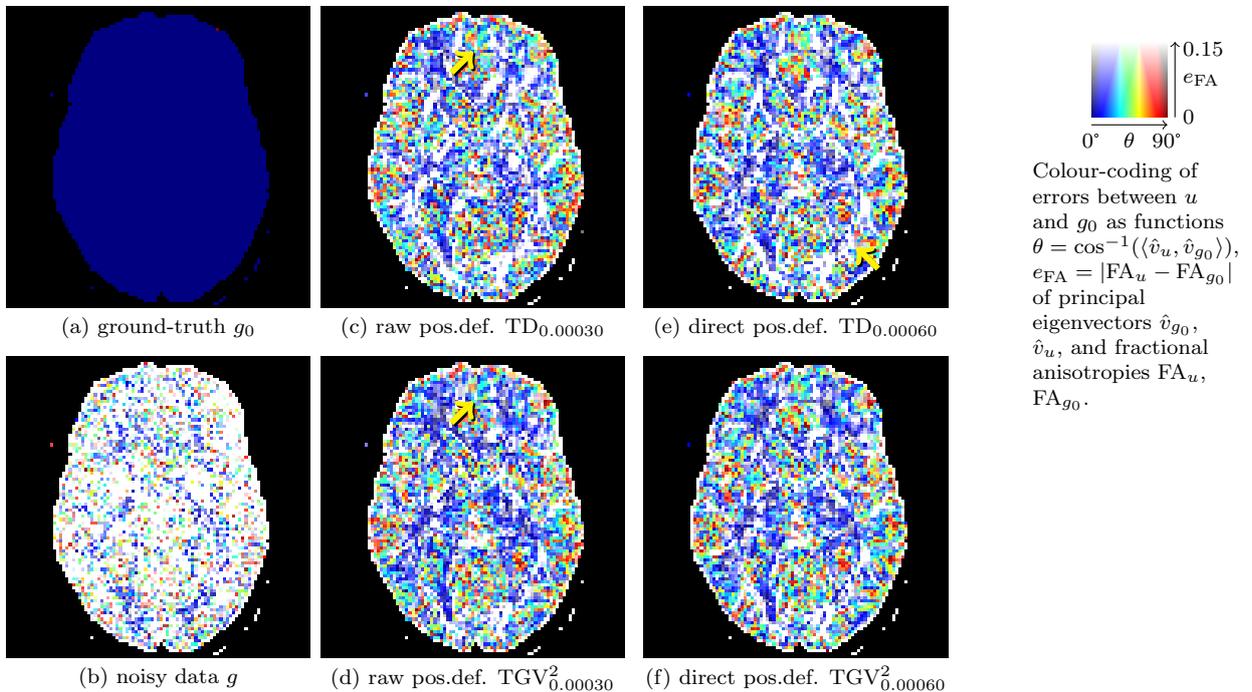


Figure 6: Colour-coded errors in the test case with additional Rician noise of parameter $\sigma = 10$ of TGV^2 and TD denoising with respect to the ground-truth g_0 , for raw and direct positivity constrained fidelity functions. Errors in fractional anisotropy appear as white, and high errors in principal eigenvectors as red. Observe in particular the high fractional anisotropy errors in (e), as indicated by the arrow, and the lesser principal eigenvector errors of TGV^2 versus TD in the area indicated by the arrow in (c) and (d).

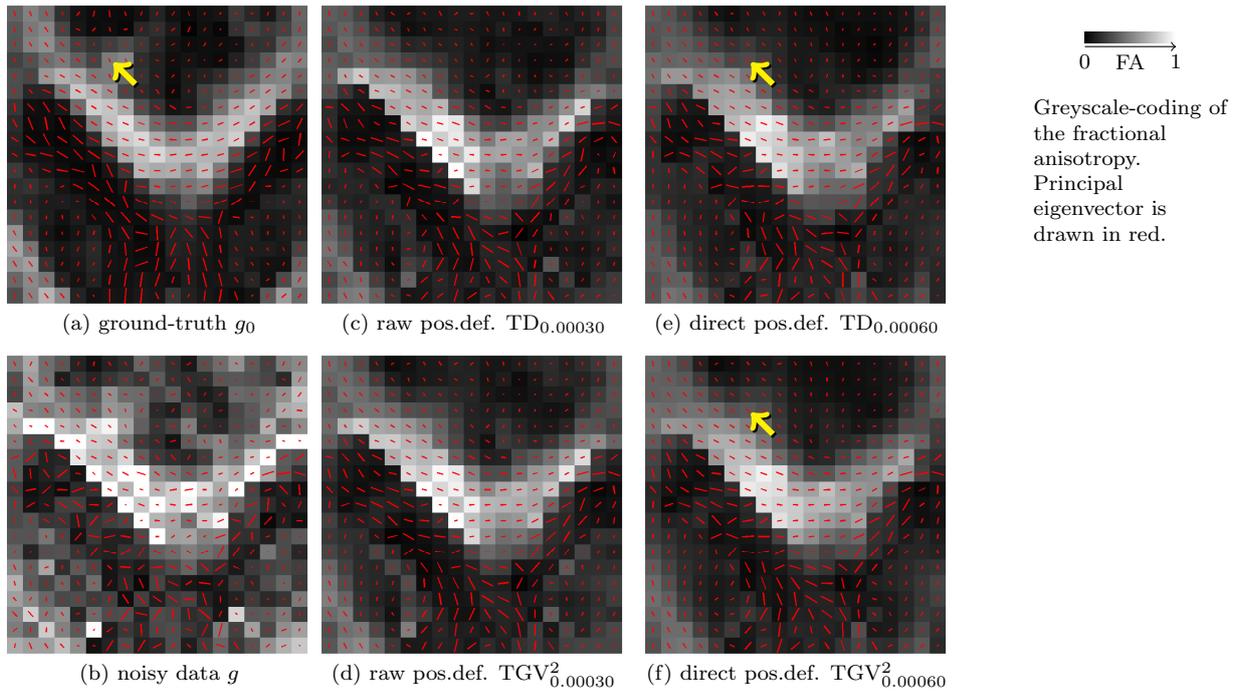


Figure 7: The *corpus callosum* region for the test case with additional Rician noise of parameter $\sigma = 10$. The fractional anisotropy is displayed in greyscale with the principal eigenvector superimposed in red. Drawn are the conventional reconstructions from full data and reduced data with additional Rician noise of parameter $\sigma = 10$ along with results of TGV^2 and TD denoising of the noisy data with the positivity constrained raw and direct fidelity functions. Observe how TGV^2 better restores the fractional anisotropy in the “protruding” area indicated by the arrow in (a), (e), and (f).