

An Optimal Transport–Based Restoration Method for Q-Ball Imaging

Thomas Vogt and Jan Lellmann

Institute of Mathematics and Image Computing (MIC), University of Lübeck
{vogt,lellmann}@mic.uni-luebeck.de

Abstract. We propose a variational approach for edge-preserving total variation (TV)-based regularization of Q-ball data from high angular resolution diffusion imaging (HARDI). While total variation is among the most popular regularizers for variational problems, its application to orientation distribution functions (ODF), as they naturally arise in Q-ball imaging, is not straightforward. We propose to use an extension that specifically takes into account the metric on the underlying orientation space. The key idea is to write the difference quotients in the TV seminorm in terms of the Wasserstein statistical distance from optimal transport. We combine this regularizer with a matching Wasserstein data fidelity term. Using the Kantorovich-Rubinstein duality, the variational model can be formulated as a convex optimization problem that can be solved using a primal-dual algorithm. We demonstrate the effectiveness of the proposed framework on real and synthetic Q-ball data.

Keywords: Variational methods, total variation, Q-ball imaging, Wasserstein distance

1 Introduction

Overview. In Q-ball imaging, one considers data $f : \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)$ given on an open, bounded, connected image domain $\Omega \subset \mathbb{R}^d$ with values in the space of *Borel probability measures* $\mathcal{P}(\mathbb{S}^2)$ over the two-dimensional unit sphere $\mathbb{S}^2 \subseteq \mathbb{R}^3$. For each $x \in \Omega$, the value $f_x := f(x)$ is a Borel probability measure on \mathbb{S}^2 . In particular, we have $f_x(A) \in [0, 1]$ for each Borel subset $A \subset \mathbb{S}^2$ and $f_x(\mathbb{S}^2) = 1$. We consider a variational approach for restoring noisy Q-ball data,

$$\inf_{u: \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)} \int_{\Omega} \delta(f_x, u_x) dx + R(u),$$

with an appropriately chosen pointwise distance $\delta(\cdot, \cdot)$ (Fig. 1). This has applications in the field of medical fiber tractography of cerebral white matter based on data that is obtained from diffusion-weighted (DW) magnetic resonance imaging (MRI) or, more precisely, high angular resolution diffusion imaging (HARDI) [24].

In this work, we focus on edge-preserving total variation (TV) regularization, i. e., $R(u)$ should encourage gradient sparsity. To this end, we adapt a specialized

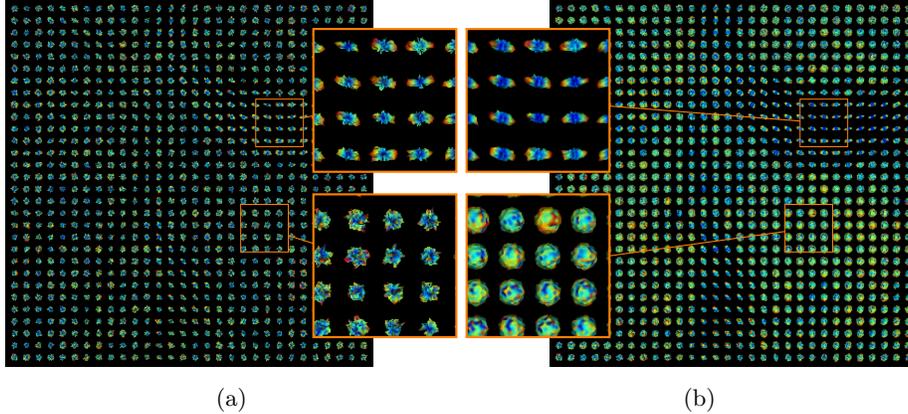


Fig. 1: Q-ball image of the corpus callosum, reconstructed from HARDI data of the human brain [21], (a) with added white Gaussian noise and (b) our TV-based reconstruction using a Wasserstein-1 data term. Noise is reduced substantially, while regions with isotropic diffusion are accurately restored or preserved.

formulation of TV that takes the metric on the space \mathbb{S}^2 into account. Similarly, for the metric $\delta(\cdot, \cdot)$, we focus on the Wasserstein-1-distance from the theory of optimal transport [25].

Motivation: Q-Ball Imaging. In medical applications, the diffusivity of water in tissues that exhibit fibrous microstructures, such as muscle fibres or axons in cerebral white matter, contains valuable information about the fiber architecture in the living organism. *Diffusion-weighted* (DW) magnetic resonance imaging (MRI) is well-established as a way of measuring the main diffusion directions by consecutively applying six or more magnetic field gradients. In *diffusion tensor imaging* (DTI) [2], few measurements are combined to extract the main diffusion direction which provides sufficient information in the case of well-aligned fiber directions. However, crossing and branching of fibers at a scale smaller than the voxel size, also called intra-voxel orientational heterogeneity (IVOH), often occurs in human cerebral white matter due to the relatively large (millimeter-scale) voxel size of DW-MRI data. Therefore, DTI data is insufficient for accurate fiber tract mapping in regions with complex fiber crossings.

High angular resolution diffusion imaging (HARDI) [23] allows for more accurate restoration of IVOH by increasing the number of applied magnetic field gradients. A widely used reconstruction scheme for HARDI data is *Q-ball imaging* [24] where the quantity of interest is the marginal probability of diffusion in a given direction, the *orientation distribution function* (ODF) [1]. However, the high angular resolution results in a larger amount of noise. Consequently, HARDI data is a particularly interesting target for postprocessing in terms of denoising and regularization [6].

Contribution. We propose a non-parametric Wasserstein-total variation regularizer for Q-ball data. Our approach is based on the observation that the regularizer in [15] amounts to a Wasserstein-based distance on the jump set. We supply the necessary theory and formally prove this connection (Sect. 2). We derive an efficient numerical formulation (Sect. 3) as a primal-dual saddle-point problem involving only convex functions. Applications to synthetic and real-world data show significant reduction of noise as well as qualitatively convincing results (Sect. 4).

Related Models. There are approaches that apply TV regularization directly to the HARDI signal $S(q)$ [17][13]. These approaches are not applicable to the ODFs of the Q-ball model which our model is aimed at. Other TV models depend on a specific parametrization of the ODFs by spherical harmonics [4][19]. In contrast, our proposed model does not depend on the specific parametrization of ODFs.

In [26], Mumford-Shah and Potts functionals are defined for manifold-valued images introducing an edge-preserving model for Q-ball data equipped with the Fisher-Rao metric following [10], [5] and [11]. However, the Fisher-Rao metric does not take the manifold structure of \mathbb{S}^2 into consideration and is not amenable to biological interpretations [18].

Our work is based on [15], where the authors derive a total variation regularizer for probability measures as a convex relaxation scheme regularization of manifold-valued images. An extension of the Kantorovich-Rubinstein formulation of the Wasserstein distance that our saddle-point formulation makes use of has been applied in [16] to the simpler problem of real-valued image denoising.

Our proposed method incorporates a TV model that does not depend on the specific parametrization of the ODFs. It is based on data fidelity and regularization terms that take the manifold structure of \mathbb{S}^2 into account, without reducing the ODF model to a simpler statistical interpretation. Furthermore, the model can be efficiently implemented using state-of-the-art primal-dual methods.

2 A W_1 - TV_{W_1} Functional for Q-Ball Data

A straightforward approach to denoising probability distributions is to apply a Rudin-Osher-Fatemi (ROF) model, based on the L^2 norm on $L^2(\mathbb{S}^2)$ and a suitable extension of the total variation (TV) regularizer, to the probability density functions:

$$\inf_{u: \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)} \int_{\Omega} \int_{\mathbb{S}^2} (f_x(z) - u_x(z))^2 dz dx + \lambda TV(u) \quad (1)$$

However, this does not allow singular measures for u_x and does not account for the (Riemannian) manifold structure of \mathbb{S}^2 : as long as the supports of u_x and f_x do not intersect, the local cost in the data term does not depend on the relative orientation of the modes (peaks) of u_x and f_x in \mathbb{S}^2 .

The Metric Space $(\mathcal{P}(\mathbb{S}^2), W_1)$. In order to account for the manifold structure of \mathbb{S}^2 , we equip $\mathcal{P}(\mathbb{S}^2)$ with the Wasserstein-1-metric

$$W_1(\mu, \mu') := \inf \left\{ \int_{\mathbb{S}^2 \times \mathbb{S}^2} d_{\mathbb{S}^2}(x, y) d\gamma(x, y) : \gamma \in \mathcal{P}(\mathbb{S}^2 \times \mathbb{S}^2), \pi_1\gamma = \mu, \pi_2\gamma = \mu' \right\}$$

that is well-known from the theory of optimal transport [25]. Here $\pi_i\gamma$ denotes the i -th marginal of the measure γ that is supported on the product space $\mathbb{S}^2 \times \mathbb{S}^2$, i.e., $\pi_1\gamma(A) := \gamma(A \times \mathbb{S}^2)$ and $\pi_2\gamma(B) := \gamma(\mathbb{S}^2 \times B)$. By $d_{\mathbb{S}^2}(\cdot, \cdot)$ we refer to the geodesic distance metric on the Riemannian manifold \mathbb{S}^2 .

Data term. Using this natural distance, we consider functionals of the form

$$\inf_{u: \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)} W_1(f_x(z), z_x(z)) + R(u).$$

If one discretizes the space of probability measures $\mathcal{P}(\mathbb{S}^2)$ using discrete probability measures on a set of points $z^1, \dots, z^l \in \mathbb{S}^2$, a naive implementation of W_1 requires to discretize the transport plan γ using l^2 points, which quickly becomes prohibitively large. Fortunately, W_1 is particularly well-suited for efficient implementation due to the Kantorovich-Rubinstein duality [12]:

$$W_1(\mu, \mu') = \sup \left\{ \int_{\mathbb{S}^2} p d(\mu - \mu') : p \in \text{Lip}_1(\mathbb{S}^2) \right\}, \quad (2)$$

where we denote by $\text{Lip}_1(\mathbb{S}^2)$ the space of Lipschitz-continuous functions on \mathbb{S}^2 with Lipschitz-constant not larger than 1.

Regularization. We claim that if one chooses the W_1 metric for the data term, the regularizer R should also be modified accordingly. In order to motivate this, we consider what the intended behavior of R should be on “cartoon-like” functions $u : \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)$ that only take two different values: let

$$u(x) = \begin{cases} u^+, & x \in U, \\ u^-, & x \in \Omega \setminus U, \end{cases} \quad (3)$$

for some fixed set $U \subset \Omega$ with smooth boundary ∂U , and $u^+, u^- \in \mathcal{P}(\mathbb{S}^2)$. Assuming for a moment that u^+ and u^- have a density, classical vector-valued TV assigns to such u a penalty of $\mathcal{H}^{d-1}(\partial U) \cdot \|u^+ - u^-\|_2$, where the $(d-1)$ -dimensional Hausdorff measure $\mathcal{H}^{d-1}(\partial U)$ is the length or area of the jump set. Similarly to the data term, we propose to replace the suboptimal 2-norm by the Wasserstein distance, i.e., we require

$$R(u) = \mathcal{H}^{d-1}(\partial U) \cdot W_1(u^+, u^-) \quad (4)$$

instead. We consider the following formulation for the regularizer R .

Definition 1. For a function $u : \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)$, we define

$$\text{TV}_{W_1}(u) := \sup \left\{ \int_{\Omega} \langle -\text{div } p(x, \cdot), u_x \rangle dx : \right. \\ \left. p \in C_c^1(\Omega \times \mathbb{S}^2; \mathbb{R}^d), p(x, \cdot) \in \text{Lip}_1(\mathbb{S}^2; \mathbb{R}^d) \right\}, \quad (5)$$

where $\langle g, \mu \rangle := \int_{\mathbb{S}^2} g(z) d\mu(z)$ whenever μ is a measure on \mathbb{S}^2 and g is a real- or vector-valued function on \mathbb{S}^2 . The divergence $\text{div } p(x, \cdot)$ is understood to be taken in the variable x only, and pointwise in the second argument.

Note that the Lipschitz constraint in (5) can be efficiently implemented as a pointwise constraint on the norm of $\nabla_z p(x, z)$. The formulation is a special case of the one proposed in [15] in the context of approximating manifold-valued variational problems. However, the authors did not include any precise results about correctness in the functional-analytic setting.

We now supply the proof that this functional has the desired property (4):

Proposition 1. Assume that U is compactly contained in Ω with C^2 -boundary ∂U . Let $u^+, u^- \in \mathcal{P}(\mathbb{S}^2)$ and let $u : \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)$ be defined as in (3). Then

$$\text{TV}_{W_1}(u) = \mathcal{H}^{d-1}(\partial U) \cdot W_1(u^+, u^-).$$

Proof. Let $p : \Omega \times \mathbb{S}^2 \rightarrow \mathbb{R}^d$ satisfy the constraints in (5) and denote by ν the outer unit normal of ∂U . The set Ω is bounded, u^+, u^- are probability measures, and $\|\text{div } p\|_{C^0(\Omega \times \mathbb{S}^2)} < \infty$ holds due to the regularity assumptions on p , therefore the following integrals converge absolutely. Using Fubini's and Gauss's theorem, we compute

$$\begin{aligned} & \int_{\Omega} \langle -\text{div } p(x, \cdot), u_x \rangle dx \\ &= \int_{\mathbb{S}^2} \int_U -\text{div } p(x, z) dx du^-(z) + \int_{\mathbb{S}^2} \int_{\Omega \setminus U} -\text{div } p(x, z) dx du^+(z) \\ &\stackrel{\text{Gauss}}{=} \int_{\mathbb{S}^2} \int_{\partial U} p(x, z) \cdot \nu(x) d\mathcal{H}^{d-1}(x) d(u^+ - u^-)(z) \\ &= \int_{\partial U} \left[\int_{\mathbb{S}^2} p(x, z) \cdot \nu(x) d(u^+ - u^-)(z) \right] d\mathcal{H}^{d-1}(x) \\ &\leq \int_{\partial U} W_1(u^+, u^-) d\mathcal{H}^{d-1}(x) = \mathcal{H}^{d-1}(\partial U) W_1(u^+, u^-). \end{aligned}$$

For the inequality, we used that $z \mapsto p(x, z)$ is 1-Lipschitz; therefore $z \mapsto p(x, z) \cdot \nu(x)$ is 1-Lipschitz as well and fulfills the constraints in the dual definition of W_1 in (2). Taking the supremum over p as in (5), we arrive at

$$\text{TV}_{W_1}(u) \leq \mathcal{H}^{d-1}(\partial U) W_1(u^+, u^-).$$

For the reverse inequality, let $\tilde{p} \in \text{Lip}_1(\mathbb{S}^2)$. By the assumption on ∂U , the unit normal field ν is continuously differentiable. Since U is also compactly

contained in Ω , there is an extension $\tilde{\nu} \in C_c^1(\Omega)$ of ν onto all of Ω . Now $p(x, z) := \tilde{p}(z)\tilde{\nu}(x)$ has the properties required in (5) and hence, by the same computation as above,

$$\begin{aligned} \text{TV}_{W_1}(u) &\geq \int_{\Omega} \langle -\text{div } p(x, \cdot), u_x \rangle dx \\ &= \int_{\partial U} \left[\int_{\mathbb{S}^2} \tilde{p}(z) d(u^+ - u^-)(z) \right] \|\nu(x)\|_2^2 d\mathcal{H}^{d-1}(x) \\ &= \mathcal{H}^{d-1}(\partial U) \int_{\mathbb{S}^2} \tilde{p}(z) d(u^+ - u^-)(z). \end{aligned}$$

Taking the supremum over all $\tilde{p} \in \text{Lip}_1(\mathbb{S}^2)$ and using the characterization (2) of W_1 shows the desired reverse inequality and concludes the proof. \square

Complete $W_1 - \text{TV}_{W_1}$ Model. Given a noisy reference image $f : \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)$, we propose to solve the variational minimization problem

$$\inf_{u: \Omega \rightarrow \mathcal{P}(\mathbb{S}^2)} \int_{\Omega} W_1(f_x, u_x) dx + \lambda \text{TV}_{W_1}(u) \quad (6)$$

using the definitions of $W_1(\cdot, \cdot)$ and $\text{TV}_{W_1}(u)$ in (2) and (5). As argued above, this naturally penalizes jumps in u by the Wasserstein distance of the left and right limit, correctly taking the metric structure of \mathbb{S}^2 into account both in the data term as well as in the regularizer.

We remark that the requirement (4) does not define the regularizer uniquely. However, as shown above, setting $R = \text{TV}_{W_1}$ satisfies the requirement, has a compact representation and can be efficiently implemented using (2).

3 Numerical Scheme

We closely follow the discretization scheme in [15] in order to formulate the problem in a saddle-point form that is amenable to standard primal-dual algorithms.

Discretization. We assume a d -dimensional image domain Ω that is discretized using n points $x^1, \dots, x^n \in \Omega$. Differentiation in Ω is done on a staggered grid with Neumann boundary conditions such that the dual operator to the differential operator D is the negative divergence with vanishing boundary values.

The model is not only valid for the manifold \mathbb{S}^2 , therefore we state the setting in full generality: We assume that the image takes values in the space of probability measures on an s -dimensional manifold \mathcal{M} that is discretized using l points $z^1, \dots, z^l \in \mathcal{M}$. If $u, v \in \mathbb{R}^{n,l}$ are the discretizations of functions on $\Omega \times \mathcal{M}$, i. e. $u_k^i \approx u(x^i, z^k)$, we define an L^2 -inner product by

$$\langle u, v \rangle_b := \sum_{i,k} b_k u_k^i v_k^i$$

using the weights vector $b \in \mathbb{R}^l$ to account for the volume element at each $z^k \in \mathcal{M}$. Gradients of functions on \mathcal{M} are defined on a staggered grid of m points $y^1, \dots, y^m \in \mathcal{M}$, such that each y^j has r neighboring points $\mathcal{N}_j \subset \{1, \dots, l\}$, $\#\mathcal{N}_j = r$, among the z^k so that the gradient $g \in \mathbb{R}^{m,s}$ of a function $p \in \mathbb{R}^l$ (i. e. $p_k \approx p(z^k)$) on the manifold \mathcal{M} is encoded by

$$A^j g^j = B^j P^j p, \text{ for each } j \in \{1, \dots, m\}.$$

The matrices $P^j \in \{0, 1\}^{r,l}$, $A^j \in \mathbb{R}^{s,s}$ and $B^j \in \mathbb{R}^{s,r}$ tie g^j to the gradient of p on the manifold \mathcal{M} ; for details we refer the reader to [15].

Discretized $W_1 - \text{TV}_{W_1}$ Model. Based on the above discretization, we can formulate saddle point forms for (6) and (1) that allow to apply a primal-dual first-order method such as [3]. In the following, the input or reference image is given by $f \in \mathbb{R}^{l,n}$ and the dimensions of the primal and dual variables are

$$u \in \mathbb{R}^{l,n}, \quad p \in \mathbb{R}^{l,d,n}, \quad g \in \mathbb{R}^{n,m,s,d}, \quad p_0 \in \mathbb{R}^{l,n}, \quad g_0 \in \mathbb{R}^{n,m,s}.$$

Note that, for $\min\{d, s\} \leq 2$ (in particular for $\mathcal{M} = \mathbb{S}^2$), explicit formulas for the orthogonal projections on the spectral norm balls are available. These projections appear in the proximal steps due to the spectral norm constraints on g (denoted by $\|\cdot\|_\sigma$).

Saddle-Point Form with Wasserstein Data Term. Using a W_1 data term, the problem's saddle point form reads

$$\begin{aligned} \min_u \max_{p,g} \quad & W_1(u, f) + \langle Du, p \rangle_b \\ \text{s.t.} \quad & u^i \geq 0, \langle u^i, b \rangle = 1, A^j g_t^{ij} = B^j P^j p_t^i, \|g^{ij}\|_\sigma \leq \lambda \forall i, j, t \end{aligned}$$

or, again applying the Kantorovich-Rubinstein duality (2) to the data term,

$$\begin{aligned} \min_u \max_{p,g,p_0,g_0} \quad & \langle u - f, p_0 \rangle_b + \langle Du, p \rangle_b \\ \text{s.t.} \quad & u^i \geq 0, \langle u^i, b \rangle = 1 \forall i, \\ & A^j g_t^{ij} = B^j P^j p_t^i, \|g^{ij}\|_\sigma \leq \lambda \forall i, j, t, \\ & A^j g_0^{ij} = B^j P^j p_0^i, \|g_0^{ij}\|_2 \leq 1 \forall i, j. \end{aligned}$$

The equality constraints can be included into the objective function by introducing suitable Lagrange multipliers.

Saddle Point Form with Quadratic Data Term. For comparison, we also implemented the quadratic model (1) using $\text{TV} = \text{TV}_{W_1}$. The quadratic data term can be implemented using the saddle point form

$$\begin{aligned} \min_u \max_{p,g} \quad & \langle u - f, u - f \rangle_b + \langle Du, p \rangle_b \\ \text{s.t.} \quad & u^i \geq 0, \langle u^i, b \rangle = 1, A^j g_t^{ij} = B^j P^j p_t^i, \|g^{ij}\|_\sigma \leq \lambda \forall i, j, t. \end{aligned}$$

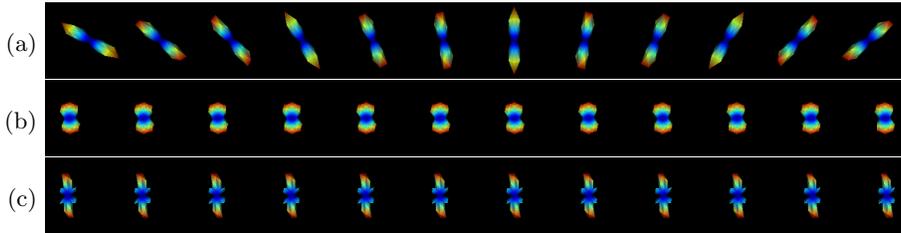


Fig. 2: (a) is a 1D image of synthetic unimodal ODFs where the main diffusion direction’s angle varies linearly from left to right. This image is strongly regularized (large λ) using (b) an L^2 data term ($\lambda = 2$) and (c) a W_1 data term ($\lambda = 10$). The L^2 data term prefers a blurred mixture of diffusion directions whereas W_1 concentrates the mass close to the median diffusion direction.

4 Experimental Results

We implemented our method in Python 3.5 using NumPy 1.11, PyCUDA 2016.1.2 and CUDA 7.5. The examples were computed on an Intel Xeon X5670 2.93GHz with 24 GB of memory and an NVIDIA GeForce GTX 480 graphics card with 1,5 GB of dedicated video memory. For each step in the primal-dual algorithm, a set of kernels was launched on the GPU, while the primal-dual gap was computed on the CPU.

Synthetic 1D Images. The synthetic images were generated using the Multi-Tensor Simulation framework `dipy.sims.voxel` included in the Dipy project [9]. The choice of generation parameters as well as the addition of complex Gaussian noise follow the description in [7].

Unimodals. A first impression of the different behavior of the W_1 data term compared to a standard L^2 data term is given in Fig. 2, where very high regularization parameters λ force the model to produce constant images. The linearly changing unimodal orientation distribution is blurred by the L^2 data term, whereas the W_1 data term tends to concentrate the mass close to the median diffusion direction. Here, the ODF is called unimodal if there is one main diffusion direction even though, due to the symmetric structure of diffusion ODFs, an unimodal ODF has two modes (on opposite sites of the sphere).

A similar experiment (Fig. 3) demonstrates that the behavior of the W_1 model is preferable if the main diffusion directions of the unimodal ODFs underlie random distortion (noise). The ground truth consists of 12 identical unimodal ODFs, while the main diffusion directions have been randomly distorted in the input image following a Gaussian distribution on the angle with 20° standard deviation.

The regularization parameter λ was chosen so that the W_1 distance to the ground truth is minimized. Both models reduce the distortion significantly: While the distances of the noisy image to the ground truth are 1.368 with respect

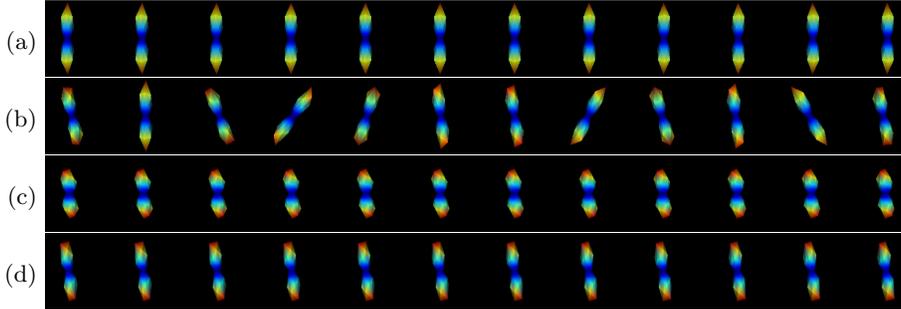


Fig. 3: Synthetic 1D Q-ball image of unimodal ODFs. (a) ground truth with constant modes and (b) noisy image with distorted diffusion directions. The noisy image was denoised using (c) a quadratic data term ($\lambda = 0.85$) and (d) the Wasserstein-1 data term ($\lambda = 2.5$). The directional noise causes blurring in the L^2 case, whereas the W_1 data term keeps the mass concentrated: The entropy of the original data is 2.267 compared to 2.336 (L^2) and 2.279 (W_1).

to W_1 , the distances of the regularized images are 0.727 (W_1 model) and 0.621 (L^2 model).

However, using the W_1 models results in improved tightness (lower entropy) of the ODFs, which is particularly promising in the context of fiber tractography, where a precise localization of the main diffusion directions is necessary for accurately tracing fine structures.

Bimodals and Edge Preservation. The next example (Fig. 4) is inspired by [26], where a similar 1D image is used for demonstration of edge preservation properties. Six voxels of the ground truth are chosen to be bimodal ODFs with the two main diffusion directions separated by 55° . The remaining six voxels are almost uniform ODFs (unimodal ODFs with almost uniform distribution of eigenvalues corresponding to the main diffusion directions). As both models use the proposed TV_{W_1} regularizer, the edge is preserved and a piecewise constant image is produced in both cases.

However, just as classical ROF models tend to reduce jump sizes across edges, and lose contrast, the result produced using the L^2 model exhibits bimodal ODFs on both sides of the jump and the tightness of the original bimodals gets lost.

Real-World Example – Human Brain HARDI Data. In order to demonstrate the applicability of our method to real-world problems, we applied it to a 2D slice from the human brain HARDI data set from [21] that contains part of the corpus callosum. To fit the HARDI data to the Q-ball ODF model, we used a spherical harmonic reconstruction model from [1] as implemented in Dipy’s `CsaOdfModel`. We added Gaussian noise following the process in [7]. For a 2D slice of 30×30 voxels, run times are appx. 28 minutes for the Wasserstein data term and 16 minutes for the quadratic data term. That’s after the relative primal-dual gap is brought down to values in the order of 10^{-6} (deviation from

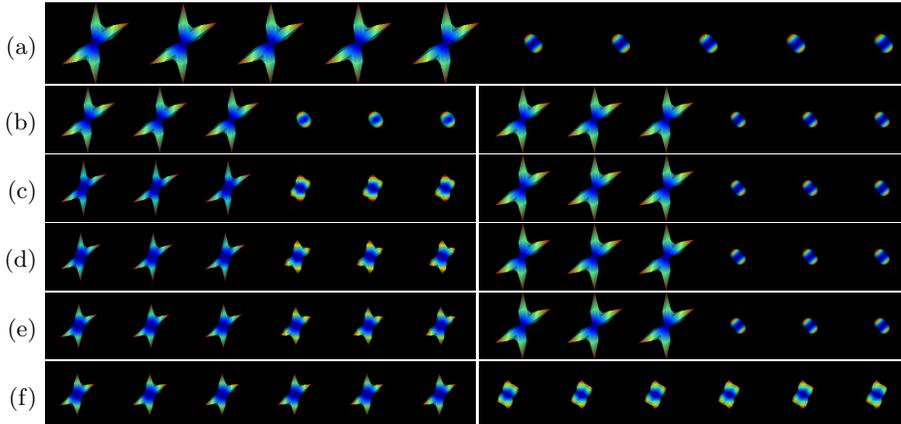


Fig. 4: Synthetic 1D Q-ball image of bimodal and almost uniform ODFs: The (a) original data was denoised (b)–(f) using an L^2 data term (left) and a W_1 data term (right) for increasing values of λ (on the left-hand side $\lambda = 0.05, 0.55, 1.05, 1.55, 2.05$ and on the right-hand side $\lambda = 0.05, 1.35, 2.65, 3.95, 5.25$). Both models preserve the edge. However, as is known from classical ROF models, the L^2 data term produces a gradual transition – i. e., contrast loss – towards the constant image, while the W_1 data term exhibits a sudden phase transition.

global minimum less than 0.001 %). The regularization parameter λ was manually chosen based on visual inspection. However, optimization of λ with respect to the distance from the ground truth leads to similar results.

Fig. 1 shows that the W_1 model succeeds at reducing the overall noise level. A comparison between the results produced by the L^2 and the W_1 model is shown in Fig. 5. While both models successfully reduce the noise, an overall blurring effect is more evident in the L^2 case – in line with our observations on the synthetic images discussed above.

5 Conclusion

We proposed a combined W_1 - TV_{W_1} variational model for restoring Q-ball data with values on the unit sphere \mathbb{S}^2 , which properly takes the metric on \mathbb{S}^2 into account, both in the data term as well as in the regularizer.

We demonstrated that our proposed model exhibits classical properties of TV regularization schemes, including preservation of edges and formation of piecewise constant regions. In connection with an L^2 data term, basic properties of ROF models, such as a loss of contrast, can be replicated. The proposed combined W_1 - TV_{W_1} model is more robust, respects the metric on the original range \mathbb{S}^2 and better preserves unimodal distributions.

The model does not require symmetry of Q-ball data on the sphere, and therefore could be easily adapted to novel asymmetric ODF approaches [8, 20].

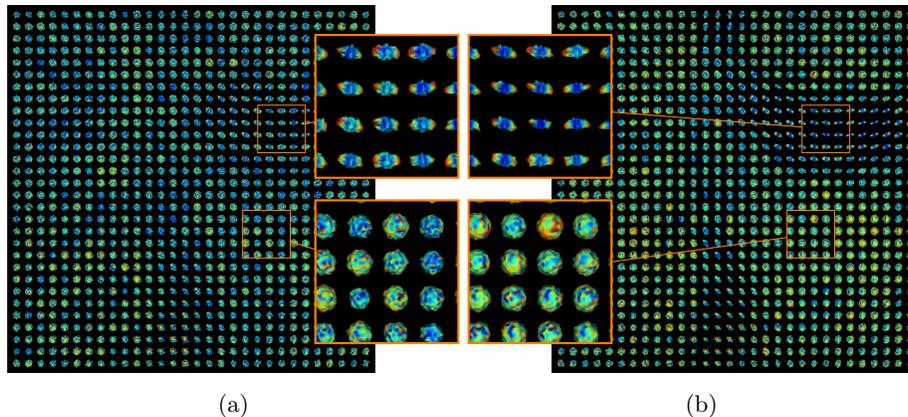


Fig. 5: Q-ball image of the corpus callosum, reconstructed from HARDI data of the human brain [21], with added white Gaussian noise and our TV-based reconstruction using (a) a quadratic data term ($\lambda = 0.11$) and (b) a Wasserstein-1 data term ($\lambda = 0.9$). The noise is reduced perceptibly in both cases, but the quadratic data term tends to blur the diffusion direction (loss of contrast).

Moreover, the approach is easily extendable to images with values in the probability space over a different manifold, or even a metric space, as they appear for example in statistical models of computer vision [22].

Despite a relatively efficient implementation using standard primal-dual algorithms, numerical performance is still limited by the high dimensionality of the problem even for small images. Developing specialized numerical methods, as well as a quantitative evaluation as a preprocessing step for fiber tractography, is subject to further work.

References

1. Aganj, I., Lenglet, C., Sapiro, G.: ODF Reconstruction in Q-Ball Imaging with Solid Angle Consideration. *Proc IEEE Int Symp Biomed Imaging* 2009. 1398-1401 (2009)
2. Basser P. J., Mattiello J., LeBihan D.: MR diffusion tensor spectroscopy and imaging. *Biophys J.* 66 (1). 259–267 (1994)
3. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* 40 (1). 120–145 (2011)
4. Chen, Y., Guo, W., Zeng, Q., Liu, Y.: A nonstandard smoothing in reconstruction of apparent diffusion coefficient profiles from diffusion weighted images. *Inverse Problems and Imaging* 2 (2). 205–224 (2008)
5. Cheng, J., Ghosh, A., Jiang, T., Deriche, R.: A Riemannian Framework for Orientation Distribution Function Computing. *Med Image Comput Comput Assist Interv* 2009 12 (1). 911–918 (2009)
6. Delputte, S., Dierckx, H., Fieremans, E., D’Asseler, Y., Achten, R., Lemahieu, I.: Postprocessing of Brain White Matter Fiber Orientation Distribution Functions. In: *ISBI '07*. 784–787 (2007)

7. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Apparent diffusion coefficients from high angular resolution diffusion imaging: estimation and applications. *Magn Reson Med* 56 (2). 395–410 (2006)
8. Ehrlicke, H. H., Otto, K. M., Klose, U.: Regularization of bending and crossing white matter fibers in MRI Q-ball fields. *Magn Reson Imaging* 29 (7). 916–926 (2011)
9. Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., Nimmo-Smith, I., Dipy Contributors: Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics* 8 (8). (2014)
10. Goh, A., Lenglet, C., Thompson, P., Vidal, R.: A nonparametric Riemannian framework for processing High Angular Resolution Diffusion Images (HARDI). In: *CVPR '09*. 2496–2503 (2009)
11. Goh, A., Lenglet, C., Thompson, P. M., Vidal, R.: A nonparametric Riemannian framework for processing high angular resolution diffusion images and its applications to ODF-based morphometry. *Neuroimage* 56 (3). 1181–1201 (2011)
12. Kantorovich, L. V., Rubinshten, G. Sh.: On a functional space and certain extremum problems. *Dokl. Akad. Nauk SSSR* 115. 1058–1061 (1957)
13. Kim, Y., Thompson, P. M., Vese, L. A.: HARDI Data Denoising Using Vectorial Total Variation and Logarithmic Barrier. *Inverse Problems and Imaging* 4 (2). 273–310 (2010)
14. Lellmann, J., Schnörr, C.: Continuous multiclass labeling approaches and algorithms. *SIAM J. Imaging Sci.* 4 (4). 1049–1096 (2011)
15. Lellmann, J., Strelakovski, E., Koetter, S., Cremers, D.: Total Variation Regularization for Functions with Values in a Manifold. In: *2013 IEEE International Conference on Computer Vision*. 2944–2951 (2013)
16. Lellmann, J., Lorenz, D. A., Schönlieb, C., Valkonen, T.: Imaging with Kantorovich-Rubinstein Discrepancy. *SIAM J. Imaging Sci.* 7 (4). 2833–2859 (2014)
17. McGraw, T., Vemuri, B., Ozarslan, E., Chen, Y., Mareci, T.: Variational denoising of diffusion weighted MRI. *Inverse Problems and Imaging* 3 (4). 625–648 (2009)
18. Ncube, S., Srivastava, A.: A novel Riemannian metric for analyzing HARDI data. *Proc. SPIE* 7962. Id. 79620Q (2011)
19. Ouyang, Y., Chen, Y., Wu, Y.: Vectorial total variation regularisation of orientation distribution functions in diffusion weighted MRI. *Int J Bioinform Res Appl* 10 (1). 110–127 (2014)
20. Reisert, M., Kellner, E., Kiselev, V.G.: About the geometry of asymmetric fiber orientation distributions. *IEEE Trans Med Imaging* 31 (6). 1240–1249 (2012)
21. Rokem, A., Yeatman, J., Pestilli, F., Wandell, B.: High angular resolution diffusion MRI. *Stanford Digital Repository*. Available at: <http://purl.stanford.edu/yx282xq2090> (2013)
22. Srivastava, A., Jermyn, I. H., Joshi, S. H.: Riemannian analysis of probability density functions with applications in vision. In: *CVPR '07*. 1–8 (2007)
23. Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., Wedeen, V. J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine* 48 (4). 577–582 (2002)
24. Tuch, D. S.: Q-ball imaging. *Magnetic Resonance in Medicine* 52 (6). 1358–1372 (2004)
25. Villani, C.: *Optimal transport. Old and new.* Grundlehren der Mathematischen Wissenschaften 338. Berlin: Springer (2009)
26. Weinmann, A., Demaret, L., Storath, M. J.: Mumford-Shah and Potts Regularization for Manifold-Valued Data. *J Math Imaging Vis* 55. 428 (2016)