

# Restoration of Manifold-Valued Images by Half-Quadratic Minimization

Ronny Bergmann\*, Raymond H. Chan†, Ralf Hielscher‡  
Johannes Persch\*, Gabriele Steidl\*

November 2, 2015

## Abstract

The paper addresses the generalization of the half-quadratic minimization method for the restoration of images having values in a complete, connected Riemannian manifold. We recall the half-quadratic minimization method using the notation of the  $c$ -transform and adapt the algorithm to our special variational setting. We prove the convergence of the method for Hadamard spaces. Extensive numerical examples for images with values on spheres, in the rotation group  $\text{SO}(3)$ , and in the manifold of positive definite matrices demonstrate the excellent performance of the algorithm. In particular, the method with  $\text{SO}(3)$ -valued data shows promising results for the restoration of images obtained from Electron Backscattered Diffraction which are of interest in material science.

## 1 Introduction

Many edge-preserving variational methods for the denoising or inpainting of real-valued images utilize the following model: let  $\mathcal{G} := \{1, \dots, n\} \times \{1, \dots, m\}$  be the image grid,  $\emptyset \neq \mathcal{V} \subseteq \mathcal{G}$  and  $\mathcal{N}(i)^+ := \{(i_1 + 1, i_2), (i_1, i_2 + 1)\}$  the set of right and upper neighbors of pixel  $i \in \mathcal{G}$ , where we suppose mirror boundary conditions. From corrupted image values  $f: \mathcal{V} \rightarrow \mathbb{R}$  we want to restore the original image  $u_0: \mathcal{G} \rightarrow \mathbb{R}$  as a minimizer of one of the following energy functionals

$$\frac{1}{2} \sum_{i \in \mathcal{V}} (f_i - u_i)^2 + \lambda \sum_{i \in \mathcal{G}} \sum_{j \in \mathcal{N}(i)^+} \varphi(|u_i - u_j|), \quad (1)$$

$$\frac{1}{2} \sum_{i \in \mathcal{V}} (f_i - u_i)^2 + \lambda \sum_{i \in \mathcal{G}} \varphi \left( \left( \sum_{j \in \mathcal{N}(i)^+} (u_i - u_j)^2 \right)^{\frac{1}{2}} \right), \quad (2)$$

where  $\lambda > 0$  is a regularization parameter and  $\varphi: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ . For  $\mathcal{V} = \mathcal{G}$  this is a typical denoising model in the presence of additive Gaussian noise. Otherwise, the model can be used for inpainting the missing image values in  $\mathcal{G} \setminus \mathcal{V}$ . Throughout this paper, we consider even functions  $\varphi: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ . For  $\varphi(t) := |t|$ , the models (1) and (2) are discrete variants of the anisotropic and isotropic Rudin-Osher-Fatemi model [42], respectively. Then the

---

\*University of Kaiserslautern, Dept. of Mathematics, Paul-Ehrlich-Str. 31, 67663 Kaiserslautern, Germany, {bergmann, persch, steidl}@mathematik.uni-kl.de

†Chinese University of Hong Kong, Dept. of Mathematics, Hong Kong, China

‡University of Chemnitz, Faculty of Mathematics, Reichenhainer Str. 39, 09107 Chemnitz, Germany, ralf.hielscher@mathematik.tu-chemnitz.de

regularization term is often referred as anisotropic/isotropic discrete total variation (TV) regularization in resemblance to its functional analytic counterpart.

**Remark 1.1.** *More generally one may consider  $\mathcal{G}$  as vertices of a graph with edge set  $\mathcal{E} := \{(i, j) : i \in \mathcal{G}, j \in \mathcal{N}(i)\}$  for some appropriate neighborhoods  $\mathcal{N}(i)$ . This is for example useful in nonlocal means approaches. Then the regularizing term sums over the edge set  $\mathcal{E}$ . For simplicity we restrict our attention to the special neighborhoods  $\mathcal{N}(i)^+$  here.*

Starting with the inaugural work [21, 22], a huge number of papers have examined the so-called half-quadratic minimization methods for solving the above restoration problems with various functions  $\varphi$  as well as other optimization problems. We only mention the ARTUR algorithm in [16]. Basically, the original problem is reformulated into an augmented one which is quadratic with respect to the image and separable with respect to additional auxiliary variables. Then an alternating minimization process is applied whose steps allow an efficient computation. Half-quadratic minimization is connected with other well-known minimization approaches. We only mention the relation to EM algorithms [14], quasi-Newton minimization [4, 33] and gradient linearization algorithms [33]. A gradient linearization method was used in particular in [53, 54] to minimize an approximate total variation regularization (2) with  $\varphi(t) := \sqrt{t^2 + \varepsilon^2}$  for  $\varepsilon \ll 1$ . It was called the “lagged diffusivity fixed point iteration” and the authors mention that it amounts to apply the multiplicative form of half-quadratic minimization to this  $\varphi$ . Finally, there is a relation to iteratively reweighted least squares methods [18, 30]. For a convergence analysis of half-quadratic minimization methods for convex functions  $\varphi$  we refer to [16, 34] and for weaker convergence results for nonconvex  $\varphi$  to [19].

In many applications signals or images having values in a manifold are of interest. Circle-valued images appear in interferometric synthetic aperture radar [13, 20] and various applications involving the phase of Fourier transformed data. Images with values in  $\mathbb{S}^2$  play a role when dealing with 3D directional information [27, 29, 51] or in the processing of color images in the chromaticity-brightness (CB) setting [15]. The motion group and the rotation group  $\text{SO}(3)$  were considered in tracking, (scene) motion analysis [38, 40, 50] and in the analysis of back scatter diffraction data [8]. Finally, images with values in the manifold of positive definite matrices appear in DT-MRI [36, 45, 55, 57] and whenever covariance matrices are adjusted to image pixels, see, e.g., [50].

Recently a TV-like model for circle-valued images was introduced in [46, 47]. For manifold-valued image restoration such an approach was proposed in [31], where the problem was reformulated as a multilabel optimization problem which was handled using convex relaxation techniques. Another method suggested in [56] employs cyclic and parallel proximal point algorithms and does not require labeling and relaxation techniques. This approach was generalized by including second order differences for circle-valued images in [9, 10], for coupled circle and real-valued images in [11] and for Riemannian manifolds in [6]. A restoration method which circumvents the direct work with manifold-valued data by embedding the matrix manifold in the appropriate Euclidean space and applying a back projection to the manifold was suggested in [41]. Recently, an iteratively reweighted least squares method for the restoration of manifold-valued images was suggested in [24]. This method can be seen as multiplicative half-quadratic minimization method for the special function  $\varphi(t) := \sqrt{t^2 + \varepsilon^2}$ .

In this paper we adopt the idea of half-quadratic minimization for general functions  $\varphi$  for the restoration of manifold-valued images. We prefer the notation of the  $c$ -transform known from optimal transport to recall the basic half-quadratic minimization approach.

Then we describe the algorithm for our problems of denoising or inpainting of manifold-valued images both in the anisotropic and isotropic case. Here we focus on the multiplicative half-quadratic minimization method. Convergence of the algorithm can be shown for images having entries in an Hadamard space. The manifold of positive definite matrices is such an Hadamard manifold. We provide several applications of the algorithm as the denoising of phase-valued images, the restoration of color images with disturbed chromaticity or of 3D directions, and the improvement of images obtained from electron backscatter diffraction of a Magnesium sample.

The outline of the paper is as follows: In Section 2 we propose our variational model and show how to handle it by the half-quadratic minimization approach. A convergence proof for the algorithm applied on Hadamard manifolds is given in Section 3. Section 4 shows various numerical examples. Finally, Appendix A contains the proofs and Appendix B lists special quantities necessary for the numerical computations.

## 2 Half-Quadratic Minimization

Let  $\mathcal{M}$  be a complete, connected  $n$ -dimensional Riemannian manifold with geodesic distance  $d: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$ . Now we consider manifold-valued images. More precisely, from corrupted image values  $f: \mathcal{V} \rightarrow \mathcal{M}$  we want to restore the original manifold-valued image  $u_0: \mathcal{G} \rightarrow \mathcal{M}$  as a minimizer of one of the following energy functionals

$$J_1(u) := \frac{1}{2} \sum_{i \in \mathcal{V}} d^2(u_i, f_i) + \lambda \sum_{i \in \mathcal{G}} \sum_{j \in \mathcal{N}(i)^+} \varphi(d(u_i, u_j)), \quad (3)$$

$$J_2(u) := \frac{1}{2} \sum_{i \in \mathcal{V}} d^2(u_i, f_i) + \lambda \sum_{i \in \mathcal{G}} \varphi \left( \left( \sum_{j \in \mathcal{N}(i)^+} d^2(u_i, u_j) \right)^{\frac{1}{2}} \right) \quad (4)$$

with  $\lambda > 0$  being a regularization parameter and  $\varphi: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ . As before we set  $\varphi(t) := \varphi(-t)$  for  $t < 0$  and consider  $\varphi$  as a function defined on the whole real axis. For  $\varphi(t) := |t|$ , the second sum in the regularizer of  $J_\nu$  is just  $\|(d(u_i, u_j))_{j \in \mathcal{N}(i)^+}\|_\nu$ ,  $\nu \in \{1, 2\}$ . Then  $J_1$  resembles the setting in [56] and is related to the anisotropic ROF functional (1) and  $J_2$  gives the approach [31] related to the isotropic case (2). In this paper we will consider smooth regularization terms, i.e., even, differentiable functions  $\varphi$ . We will compute a minimizer of  $J_\nu$ ,  $\nu \in \{1, 2\}$ , by half-quadratic minimization methods.

In the following, we briefly recall the reformulation idea of half-quadratic minimization using the concept of the  $c$ -transform and apply it to (3) and (4). The  $c$ -transform of functions defined on metric spaces is used in connection with optimal transport problems see, e.g., [52, p. 86f] and seems also to be an appropriate approach here. Given a function  $c: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , the  $c$ -transform of a function  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  is defined by

$$\varphi^c(s) := \inf_{t \in \mathbb{R}} \{c(t, s) - \varphi(t)\}.$$

By this definition we see that  $\varphi(t) + \varphi^c(s) \leq c(t, s)$ . For  $c(t, s) := -st$  we have  $\varphi^c = -(-\varphi)^*$  with the Fenchel transform defined for  $h: \mathbb{R} \rightarrow \mathbb{R}$  as

$$h^*(s) := \sup_{t \in \mathbb{R}} \{ts - h(t)\}.$$

Recall that  $h^{**} = h$  if and only if  $h$  is lower semi-continuous (lsc) and convex. In the

*multiplicative* and *additive* half-quadratic methods we use the functions

$$c(t, s) := t^2 s, \quad (\text{multiplicative}) \quad (5)$$

$$c(t, s) := \frac{1}{2} \left( \sqrt{a} t - \frac{1}{\sqrt{a}} s \right)^2, \quad a > 0, \quad (\text{additive}) \quad (6)$$

respectively. The multiplicative setting was introduced in [21] and the additive one in [22]. The quadratic cost function  $c(t, s)$  in the additive setting is also handled in optimal transport topics. Note that the cost function  $c(t, s)$  in the multiplicative case is not bounded from below in  $s$ . The following proposition is crucial for the half-quadratic reformulation.

**Proposition 2.1.** *Let  $\varphi: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  be an even, differentiable function and let  $c: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  be given by (5) and (6), respectively.*

i) *If the functions*

$$\begin{aligned} \Phi(t) &:= \begin{cases} -\varphi(\sqrt{t}) & \text{for } t \geq 0, \\ +\infty & \text{for } t < 0, \end{cases} & (\text{multiplicative}) \\ \Phi(t) &:= \frac{1}{2} a t^2 - \varphi(t), & (\text{additive}) \end{aligned}$$

*respectively, are convex, then  $\varphi = \varphi^{cc}$ , i.e., setting  $\psi(s) := \varphi^c(s)$  one has*

$$\psi(s) = \inf_{t \in \mathbb{R}} \{c(t, s) - \varphi(t)\}, \quad (7)$$

$$\varphi(t) = \inf_{s \in \mathbb{R}} \{c(t, s) - \psi(s)\}. \quad (8)$$

ii) *If in addition to the assumption in i) we have*

$$\lim_{t \rightarrow \infty} \frac{\varphi(t)}{t^2} \rightarrow 0, \quad (\text{multiplicative}) \quad (9)$$

$$\lim_{t \rightarrow \infty} \frac{\varphi(t)}{t^2} < \frac{1}{2} a, \quad (\text{additive}) \quad (10)$$

*respectively, and in the multiplicative case also  $\varphi'(t) \geq 0$  for  $t \geq 0$  and  $\varphi''(0+) := \lim_{t \rightarrow 0+} \frac{\varphi'(t)}{t}$  exists, then the infimum in (7) and (8) is attained for  $(t, s) = (t, s(t))$  with*

$$\begin{aligned} s(t) &:= \begin{cases} \frac{\varphi'(t)}{2t} & \text{for } t > 0, \\ \frac{\varphi''(0+)}{2} & \text{for } t = 0, \end{cases} & (\text{multiplicative}) \\ s(t) &:= a t - \varphi'(t), & (\text{additive}) \end{aligned} \quad (11)$$

*respectively, and for these pairs we have  $\varphi(t) + \psi(s) = c(t, s)$ . The choice is unique except for the multiplicative case and  $t = 0$ , where any  $s$  larger than  $\frac{\varphi''(0+)}{2}$  is also a solution.*

iii) *If in the multiplicative case in addition  $\varphi'(t) > 0$  for  $t > 0$  and  $\varphi''(0+) > 0$ , then  $s \in (0, \frac{\varphi''(0+)}{2}]$ .*

Note that in the multiplicative case  $\psi(s) = -\infty$  for  $s < 0$ , so that we can restrict our attention in the infimum in (8) to  $s \geq 0$ . Further, the assumption  $\varphi'(t) \geq 0$ ,  $t \geq 0$ , is in particular fulfilled if  $\varphi(t)$  is convex and  $\varphi''(0+) \geq 0$ . The proof can be given following for

example the lines in [16, 34]. However, since the assumptions in these papers are slightly different and the  $c$ -transform notation is not used, we add the proof in the appendix A to make the paper self-contained. The functions  $\varphi$  which fulfill the conditions in Proposition 2.1 and which were used in our numerical test are listed in Table 1. Further examples are collected in [16, 33, 34].

In the following we assume that  $\varphi$  fulfills the assumptions of Proposition 2.1. Now the idea is to replace  $\varphi$  in (3), resp., (4) by the expression in (8) and to consider  $\min_u J_\nu(u) = \min_u \min_v \mathcal{J}_\nu(u, v)$  with

$$\mathcal{J}_1(u, v) := \frac{1}{2} \sum_{i \in \mathcal{V}} d^2(u_i, f_i) + \lambda \sum_{i \in \mathcal{G}} \sum_{j \in \mathcal{N}(i)^+} \left( c(d(u_i, u_j), v_{i,j}) - \psi(v_{i,j}) \right), \quad (12)$$

$$\mathcal{J}_2(u, v) := \frac{1}{2} \sum_{i \in \mathcal{V}} d^2(u_i, f_i) + \lambda \sum_{i \in \mathcal{G}} \left( c \left( \left( \sum_{j \in \mathcal{N}(i)^+} d^2(u_i, u_j) \right)^{\frac{1}{2}}, v_i \right) - \psi(v_i) \right), \quad (13)$$

where we have used the notation  $v := (v_{i,j})_{i,j \in \mathcal{G}}$  in the anisotropic case and  $v := (v_i)_{i \in \mathcal{G}}$  in the isotropic case. By Proposition 2.1 i), minimizing  $\mathcal{J}_\nu$  over  $u$  and  $v$  gives the same solutions for  $u$  as just minimizing  $J_\nu$  over  $u$ . More precisely we give the following remark.

**Remark 2.2.** Let us abbreviate  $d_{i,j} := d(u_i, u_j)$ ,  $d_u := (d_{i,j})_{i,j \in \mathcal{G}}$  in the anisotropic case and  $d_i := \left( \sum_{j \in \mathcal{N}(i)^+} d^2(u_i, u_j) \right)^{\frac{1}{2}}$ ,  $d_u := (d_i)_{i \in \mathcal{G}}$  in the isotropic case. If  $\hat{u}$  is a minimizer of  $J_\nu$ ,  $\nu = 1, 2$ , then  $(\hat{u}, s(d_{\hat{u}}))$  is a minimizer of  $\mathcal{J}_\nu$ ,  $\nu = 1, 2$  and conversely. In particular, if

$$\hat{u} = \arg \min_u \mathcal{J}_\nu(u, s(\hat{u}))$$

holds true, then  $\hat{u}$  is a minimizer of  $J_\nu$ .

Now we can apply an alternating minimization over  $v_{i,j} \in \mathbb{R}$ , resp.,  $v_i \in \mathbb{R}$  and  $u \in \mathcal{M}^{n \times m} =: M$ :

$$v^{(k+1)} \in \arg \min_v \mathcal{J}_\nu(u^{(k)}, v), \quad (14)$$

$$u^{(k+1)} \in \arg \min_u \mathcal{J}_\nu(u, v^{(k+1)}). \quad (15)$$

Clearly, under the assumptions of Proposition 2.1 we have

$$\mathcal{J}_\nu(u^{(k)}, v^{(k+1)}) = J_\nu(u^{(k)}). \quad (16)$$

We want to work with the differentiable function  $d^2$  in the second iteration (15). Since the additive reformulation leads to a non-differentiable function  $d^2 + sd$ , we restrict our attention in the following to the multiplicative case.

**Minimization with respect to  $v$ .** The minimization over  $v$  in (14) can be done separately for the  $v_{i,j}$  or  $v_i$ . By Proposition 2.1 a minimizer is given by

$$v^{(k+1)} := s(d_{u^{(k)}}),$$

where  $s$  is defined as in (11). Note that only for  $d(u_i^{(k)}, u_j^{(k)}) = 0$  in the anisotropic case and  $\sum_{j \in \mathcal{N}(i)^+} (d^2(u_i^{(k)}, u_j^{(k)}))^{\frac{1}{2}} = 0$  in the isotropic case a larger value also could be taken as a minimizer.

If  $\varphi$  fulfills the assumptions of Proposition 2.1 iii), then  $v^{(k+1)} \in (0, \varphi''(0+)/2]$ .

**Minimization with respect to  $u$ .** The minimization over  $u$  in (15) is equivalent to finding the minimizer of

$$\begin{aligned}\mathcal{J}_{1,v^{(k)}}(u) &:= \frac{1}{2} \sum_{i \in \mathcal{V}} d^2(u_i, f_i) + \lambda \sum_{i \in \mathcal{G}} \sum_{j \in \mathcal{N}(i)^+} d^2(u_i, u_j) v_{i,j}^{(k)}, \\ \mathcal{J}_{2,v^{(k)}}(u) &:= \frac{1}{2} \sum_{i \in \mathcal{V}} d^2(u_i, f_i) + \lambda \sum_{i \in \mathcal{G}} \left( \sum_{j \in \mathcal{N}(i)^+} d^2(u_i, u_j) \right) v_i^{(k)},\end{aligned}$$

respectively. We can apply, e.g., a gradient descent or a Riemann–Newton method, see [1]. Both methods are described in the following for our setting and were implemented.

We need the following notation. Let  $T_x \mathcal{M}$  denote the tangential space of  $\mathcal{M}$  at  $x \in \mathcal{M}$  and  $\langle \cdot, \cdot \rangle_x : T_x \mathcal{M} \times T_x \mathcal{M} \rightarrow \mathbb{R}$  the Riemannian metric with induced norm  $\|\cdot\|_x$ . Let  $\gamma_{x,\xi}(t)$ ,  $x \in \mathcal{M}$ ,  $\xi \in T_x \mathcal{M}$  be the minimal geodesic starting from  $\gamma_{x,\xi}(0) = x$  with  $\dot{\gamma}_{x,\xi}(0) = \xi$ . Then the exponential map  $\exp_x : T_x \mathcal{M} \rightarrow \mathcal{M}$  is given by  $\exp_x \xi = \gamma_{x,\xi}(1)$ . The inverse exponential map denoted by  $\log_x = \exp_x^{-1} : \mathcal{M} \rightarrow T_x \mathcal{M}$  is locally well-defined. For the manifolds used in our numerical examples, namely the sphere  $\mathbb{S}^n$ ,  $n \in \mathbb{N}$ , the  $\text{SO}(3)$  and the manifold  $\mathcal{P}(r)$ ,  $r \in \mathbb{N}$ , of symmetric positive definite  $r \times r$  matrices, the specific maps are given in the Appendix B. Finally, for  $F : \mathcal{M} \rightarrow \mathbb{R}$ , let

$$\text{grad } F(x) \in T_x \mathcal{M} \quad \text{and} \quad \text{Hess}_F(x) : T_x \mathcal{M} \rightarrow T_x \mathcal{M}$$

be the Riemannian gradient and the Hessian of  $F : \mathcal{M} \rightarrow \mathbb{R}$  at  $x \in \mathcal{M}$ , respectively. For  $d(\cdot, y) : \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$  one has

$$\text{grad } d^2(x, y) = -2 \log_x y.$$

Considering the image  $u \in \mathcal{M}$ , we abbreviate the gradient and the Hessian of a function  $F : \mathcal{M} \rightarrow \mathbb{R}$  defined on the product manifold  $\mathcal{M}$  also by  $\text{grad } F$  and  $\text{Hess}_F$ , resp., since its use becomes clear from the context.

**Gradient descent method.** The gradient descent method computes, starting with  $\tilde{u}^{(0)} := u^{(k)}$ , iteratively

$$\tilde{u}^{(r+1)} = \exp_{\tilde{u}^{(r)}} \left( t_r \text{grad } \mathcal{J}_{\nu, v^{(k+1)}}(u) \Big|_{u=\tilde{u}^{(r)}} \right), \quad \nu \in \{1, 2\}, \quad (17)$$

with appropriate step sizes  $t_r > 0$ . The gradient  $\text{grad } \mathcal{J}_{\nu, v^{(k+1)}}$  is given by

$$\left( \text{grad } \mathcal{J}_{1,v^{(k+1)}}(u) \right)_i = -1_{\mathcal{V}(i)} \log_{u_i} f_i - 2\lambda \sum_{j \in \mathcal{N}(i)^+} v_{i,j}^{(k+1)} \log_{u_i} u_j, \quad (18)$$

$$\left( \text{grad } \mathcal{J}_{2,v^{(k+1)}}(u) \right)_i = -1_{\mathcal{V}(i)} \log_{u_i} f_i - 2\lambda \left( v_i^{(k+1)} \sum_{j \in \mathcal{N}(i)^+} \log_{u_i} u_j - \sum_{j \in \mathcal{N}(i)^-} v_j^{(k+1)} \log_{u_i} u_j \right) \quad (19)$$

where  $i \in \mathcal{G}$ ,  $\mathcal{N}(i)^- := \{(i_1 - 1, i_2), (i_1, i_2 - 1)\}$  and  $\mathcal{N}(i) := \mathcal{N}(i)^+ \cup \mathcal{N}(i)^-$ .

**Riemann–Newton method.** Alternatively we can use a Riemann–Newton method to compute a minimizer. Finding a descent direction  $\eta_r \in T_{\tilde{u}^{(r)}} \mathcal{M}$  with Newton's method is done for  $\nu \in \{1, 2\}$  by solving the system of equations

$$\text{Hess}_{\mathcal{J}_{\nu, v^{(k+1)}}}(\tilde{u}^{(r)})(\eta_r) = -\text{grad } \mathcal{J}_{\nu, v^{(k+1)}}(\tilde{u}^{(r)}). \quad (20)$$

---

**Algorithm 1** Image Restoration by Half-Quadratic Minimization (multiplicative)

---

**Input:**  $\mathcal{V}$ , corrupted image  $f \in \mathcal{M}^{\#\mathcal{V}} \subseteq \mathcal{M}$ ,  $\lambda, \varphi$

**Output:** Restored image  $u \in \mathcal{M}$

**Initialize**  $u^{(0)}$

**repeat**

$k \leftarrow k + 1;$

$v^{(k+1)} = s(d^{(k)})$ ;

Compute

$u^{(k+1)} := \exp_{u^{(R)}} \eta_R$

by  $R$  steps of a gradient descent method (18), resp. (19)

or by Newton's approach (20);

**until** stopping criterion is reached;

---

Then we update, starting with  $\tilde{u}^{(0)} := u^{(k)}$ , iteratively

$$\tilde{u}^{(r+1)} = \begin{cases} \exp_{\tilde{u}^{(r)}} \eta_r & \text{if } \langle \eta_r, \text{grad } \mathcal{J}_{\nu, v^{(k+1)}}(\tilde{u}^{(r)}) \rangle_{\tilde{u}^{(r)}} < 0, \\ \exp_{\tilde{u}^{(r)}}(-\text{grad } \mathcal{J}_{\nu, v^{(k+1)}}(\tilde{u}^{(r)})) & \text{otherwise.} \end{cases}$$

The whole half-quadratic minimization method for our problem is given in Algorithm 1.

In [33] it was shown that the multiplicative half-quadratic minimization is equivalent to the quasi-Newton descent method, and therefore we expect that its performance is better than the simple gradient descent method. Let us comment on this for the manifold-valued setting.

**Remark 2.3.** (*Relation of half-quadratic minimization to gradient descent and (quasi) Newton methods*)

We restrict our attention to the case  $\nu = 1$ . Similar considerations can be done for  $\nu = 2$ . The gradient of the initial functional  $J_1$  in (3) is given for  $u_i \neq u_j$  by

$$\begin{aligned} (\text{grad } J_1(u))_i &= -1_{\mathcal{V}(i)} \log_{u_i} f_i - \lambda \sum_{j \in \mathcal{N}(i)} \varphi'(d(u_i, u_j)) \frac{\log_{u_i} u_j}{\|\log_{u_i} u_j\|_{u_i}} \\ &= -1_{\mathcal{V}(i)} \log_{u_i} f_i - 2\lambda \sum_{j \in \mathcal{N}(i)} \frac{\varphi'(d(u_i, u_j))}{2d(u_i, u_j)} \log_{u_i} u_j \end{aligned}$$

which by (11) can be rewritten as

$$(\text{grad } J_1(u))_i = -1_{\mathcal{V}(i)} \log_{u_i} f_i - 2\lambda \sum_{j \in \mathcal{N}(i)} s(d(u_i, u_j)) \log_{u_i} u_j. \quad (21)$$

Hence a gradient descent algorithm applied to the initial functional  $J_1$  coincides with the half-quadratic method if we perform only one step in the gradient descent method (17) to obtain an update of  $u$ . More than one gradient descent step in (17) leads to a linearized gradient descent of  $J_1$ . If we perform a Riemann–Newton step (20) to update  $u$  we have a quasi-Newton method for  $J_1$ . Note that  $v^{(k+1)} = s(d_{u^{(k)}})$  is fixed in the half-quadratic update step of  $u$  which is not the case in (21). This simplifies the computation of the Hessian in the half-quadratic approach.

### 3 Convergence in Hadamard Manifolds

We start with a general remark.

**Remark 3.1.** Assume that  $\varphi: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  fulfills the assumptions of Proposition 2.1 iii). Let  $\{(u^{(k)}, v^{(k)})\}_k$  be the sequence produced by Algorithm 1. Then we know that  $v^{(k)} \in (0, \varphi''(0+)/2]^p$ , where  $p = 2\#\mathcal{G}$ , if  $\nu = 1$  (anisotropic case), and  $p = \#\mathcal{G}$ , if  $\nu = 2$  (isotropic case). By construction we have for the iterates produced by Algorithm 1 that

$$\mathcal{J}_\nu(u^{(k)}, v^{(k)}) \geq \mathcal{J}_\nu(u^{(k)}, v^{(k+1)}) \geq \mathcal{J}_\nu(u^{(k+1)}, v^{(k+1)}) \quad (22)$$

so that the sequence  $\{\mathcal{J}_\nu(u^{(k)}, v^{(k)})\}_{k \in \mathbb{N}}$  is monotonically decreasing. By (8) and since  $\varphi$  is nonnegative, the function  $\mathcal{J}_\nu$  is bounded from below by zero and the sequence  $\{\mathcal{J}_\nu(u^{(k)}, v^{(k)})\}_k$  converges to some  $b_\nu$ . This holds also true for  $J_\nu(u^{(k)})$  by (16). If  $\mathcal{M}$  is compact as in the case of spheres or  $\text{SO}(3)$ , then  $\{u^{(k)}\}_k$  is clearly bounded. If  $\mathcal{M}$  is an Hadamard space as defined in the next subsection and  $\varphi$  is coercive, then  $\{u^{(k)}\}_k$  is bounded since  $J_\nu$  is by Proposition 3.3 coercive. In these cases  $\{(u^{(k)}, v^{(k)})\}_k$  is also bounded and therefore there exists a subsequence  $\{(u^{(k_j)}, v^{(k_j)})\}_j$  which converges to a point  $(\bar{u}, \bar{v}) \in \mathcal{M} \times [0, \varphi''(0)/2]^p$ .

For Hadamard spaces many results on the convergence of algorithms carry directly over from the Hilbert space setting. This is in particular true for the half-quadratic minimization algorithm. In this section we summarize these results for convex functions  $\varphi$  and data in Hadamard spaces.

We start by recalling some basic facts. A curve  $\gamma: [0, 1] \rightarrow X$  in a metric space  $(X, d)$  is called a geodesic if for all  $t_1, t_2 \in [0, 1]$  the relation

$$d(\gamma(t_1), \gamma(t_2)) = |t_1 - t_2|d(\gamma(0), \gamma(1))$$

holds true. A function  $h: X \rightarrow \mathbb{R}$  is called convex if  $h \circ \gamma$  is convex for each geodesic  $\gamma: [0, 1] \rightarrow X$ , i.e., if for all  $t \in [0, 1]$  we have

$$h(\gamma(t)) \leq th(\gamma(0)) + (1 - t)h(\gamma(1))$$

and strictly convex if we have a strict inequality for all  $0 < t < 1$ . An Hadamard space is a complete metric space  $(\mathcal{H}, d)$  with the property that any two points  $x, y$  are connected by a geodesic and the following condition holds true

$$d(x, v)^2 + d(y, w)^2 \leq d(x, w)^2 + d(y, v)^2 + 2d(x, y)d(v, w), \quad (23)$$

for any  $x, y, v, w \in X$ . Inequality (23) implies that Hadamard spaces have nonpositive curvature [3, 39] and Hadamard spaces are thus a natural generalization of complete simply connected Riemannian manifolds of nonpositive sectional curvature, the so-called Hadamard manifolds. For more details, the reader is referred to [5, 26]. Unfortunately, the spheres and the rotation group are not Hadamard manifolds, while the symmetric positive definite matrices have this nice property. The following facts can be shown similarly as in  $\mathbb{R}^d$ , see [5, Lemma 2.2.9], [48].

**Lemma 3.2.** Let  $(\mathcal{H}, d)$  be an Hadamard space and  $F: \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$  be a convex lsc function which is coercive, i.e., satisfies  $F(x) \rightarrow +\infty$  whenever  $d(x, x_0) \rightarrow +\infty$  for some  $x_0 \in \mathcal{H}$ . Then  $F$  has a minimizer. If  $F$  is convex, then any critical point is a global minimizer. If  $F$  is coercive and strictly convex, then the minimizer is unique.

In an Hadamard space  $(\mathcal{H}, d)$  we have that

- (D1)  $d: \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}_{\geq 0}$  and  $d^2: \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}_{\geq 0}$  are convex, and
- (D2)  $d^2(\cdot, y): \mathcal{H} \rightarrow \mathbb{R}_{\geq 0}$  is strictly convex.

Then we obtain the following proposition whose simple proof is added for convenience in Appendix A.

**Proposition 3.3.** *Let  $(\mathcal{M}, d) = (\mathcal{H}, d)$  be an Hadamard manifold.*

- i) *Let  $\varphi: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  and in the case  $\mathcal{V} \neq \mathcal{G}$ , further assume that  $\varphi$  is coercive. Then the functions  $J_\nu$ ,  $\nu = 1, 2$ , in (3) and (4) are coercive so that they have a minimizer.*
- ii) *If in addition  $\varphi$  is increasing and convex, then the functions  $J_\nu$ ,  $\nu = 1, 2$ , are convex. If in addition  $\mathcal{V} = \mathcal{G}$  or  $\varphi$  is strictly convex, then the functions  $J_\nu$ ,  $\nu = 1, 2$ , are strictly convex and have unique minimizers.*

Under the assumptions of Proposition 2.1 iii), we have in our algorithm that  $v^{(k)} > 0$ . Then, we see similarly as in the proof of Proposition 3.3 that the functionals  $J_{\nu, v^{(k)}}$ ,  $\nu = 1, 2$ , are coercive and strictly convex. Thus the minimizer  $u^{(k)}$  exists and is unique.

**Theorem 3.4.** *Let  $(\mathcal{M}, d) = (\mathcal{H}, d)$  be an Hadamard manifold. Let  $\varphi: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  be an even, continuously differentiable, convex function which fulfills*

- i)  $\varphi(\sqrt{t})$ ,  $t > 0$  is concave,
- ii)  $\lim_{t \rightarrow \infty} \frac{\varphi(t)}{t^2} \rightarrow 0$ ,
- iii)  $\varphi''(0+) > 0$ .

*In the case  $\mathcal{V} \neq \mathcal{G}$  we further assume that  $\varphi$  is strictly convex. Then the sequence  $\{u^{(k)}\}_{k \in \mathbb{N}}$  generated by Algorithm 1 converges to the minimizer of  $J_\nu$ ,  $\nu = 1, 2$ .*

The proof which follows standard arguments is given in the appendix A. Note that the assumptions of the theorem are fulfilled for the first two functions in Table 1.

**Remark 3.5.** *The assumptions in the theorem include those in Proposition 2.1. Note that the convexity of  $\varphi$  and  $\varphi''(0+) > 0$  imply that  $\varphi'(t) > 0$  for all  $t > 0$ . Additionally the continuity of  $\varphi'$  is required to make the function  $s$  in (11) continuous and the (strict) convexity of  $\varphi$  to make the objective function strictly convex. Since  $\varphi$  is convex, its derivative is increasing. Together with  $\varphi''(0+) > 0$  this implies that  $\varphi'(t) > 0$  for all  $t > 0$  so that  $\varphi$  is increasing for  $t > 0$  and coercive.*

## 4 Numerical Examples

In this section we demonstrate the performance of Algorithm 1 for the functions  $\varphi$  from Table 1. These functions are known for their edge-preserving properties. Note that  $\varphi_1$  was used in the ‘lagged diffusivity fixed point iteration’ [53] for real-valued images and in the iteratively re-weighted least squares method [24] for  $\mathbb{S}^2$ -valued and  $\mathcal{P}(3)$ -valued images. The function  $\varphi_2$  is a Moreau envelope of the absolute value function, also known as the Huber function. Both functions are convex. The non-convex function  $\varphi_3$  was used for edge-preserving restoration of real-valued images in [16, 33].

Unless stated otherwise we use the anisotropic approach and Newton’s method in our implementations. Although neither the spheres nor the rotation group are Hadamard manifolds we have observed convergence in all our numerical examples. This may be

|                | $\varphi(t)$  | $s(t)$   |
|----------------|---|--|
| $\varphi_1(t)$ | $\sqrt{t^2 + \varepsilon^2}$  | $\frac{1}{\sqrt{t^2 + \varepsilon^2}}$   |
| $\varphi_2(t)$ | $\begin{cases} \frac{1}{2}t^2 & t < \varepsilon, \\ \varepsilon t  - \frac{1}{2}\varepsilon^2 & t \leq \varepsilon \end{cases}$ | $\begin{cases} 1 & t < \varepsilon, \\ \frac{\varepsilon}{ t } & t \leq \varepsilon \end{cases}$ |
| $\varphi_3(t)$ | $1 - \exp(-\varepsilon^2 t^2)$  | $\varepsilon^2 \exp(-\varepsilon^2 t^2)$   |

**Table 1.** Functions  $\varphi$  fulfilling the assumptions of Proposition 2.1.

due to the fact that neighboring image pixels have values which are close enough on the manifold.

The algorithms were implemented in MATLAB Version 14b. The computations were performed on a Dell with 8 GB of RAM and an Intel Core i7, 2.93 GHz, on Ubuntu 14.04 LTS.

#### 4.1 $\mathbb{S}^1$ -valued data

We start with the one-dimensional signal in Fig. 1 to show how the different functions  $\varphi$  from Table 1 perform and how the parameter  $\varepsilon$  influences the results. The original signal in Fig. 1 (a) was obtained from  $f(x) = 8\pi x^2$  by sampling with size 0.01 and unwrapping modulo  $2\pi$  such that the data are represented in  $[-\pi, \pi]$ . Then wrapped Gaussian noise with standard deviation  $\sigma = 0.3$  was added. Using  $\varphi_1$  to restore the signal gives relatively large error in Fig. 1 (b). The Huber function  $\varphi_2$  and the exponential function  $\varphi_3$  show better results in Fig. 1 (c) and 1 (d). The regularization parameter  $\lambda$  was adapted to get the best error

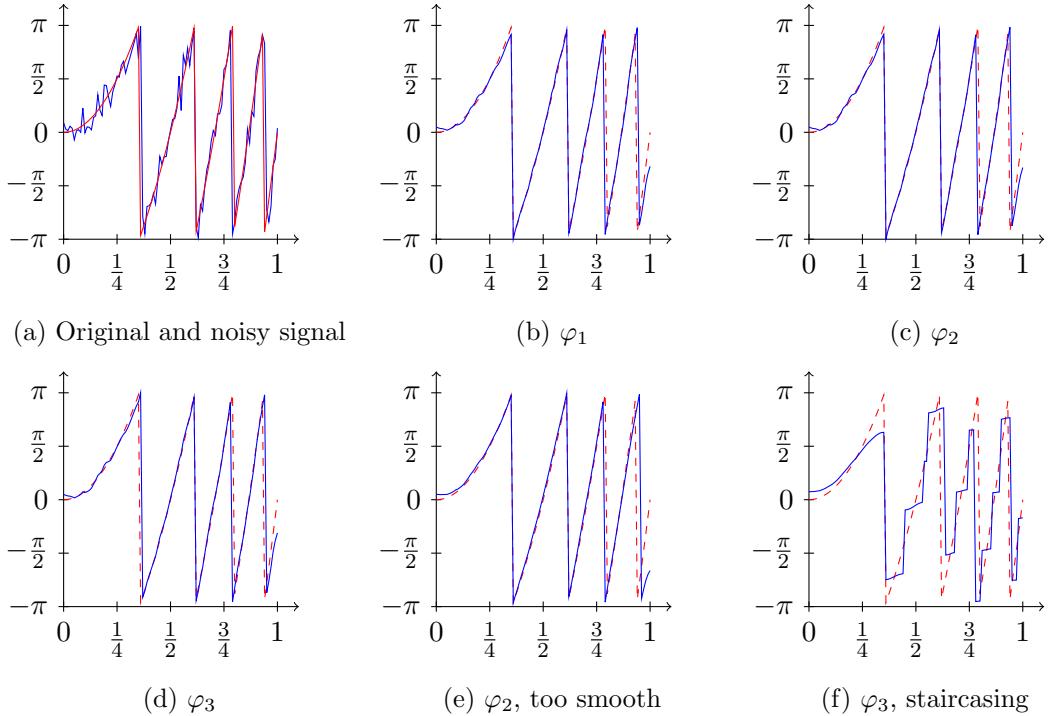
$$\text{err} := \frac{1}{N} \sum_{i=0}^{N-1} d(f_i, u_i),$$

where  $N = 101$ . Making  $\varepsilon$  in the Huber function larger leads to the smoother result in Fig. 1 (e) which approximates the original signal only well at the beginning of the signal. In Fig. 1 (f) we choose a larger  $\varepsilon$  in the function  $\varphi_3$  with the effect that edges of smaller height are smoothed and we have a staircasing effect for nearly equally high ascents.

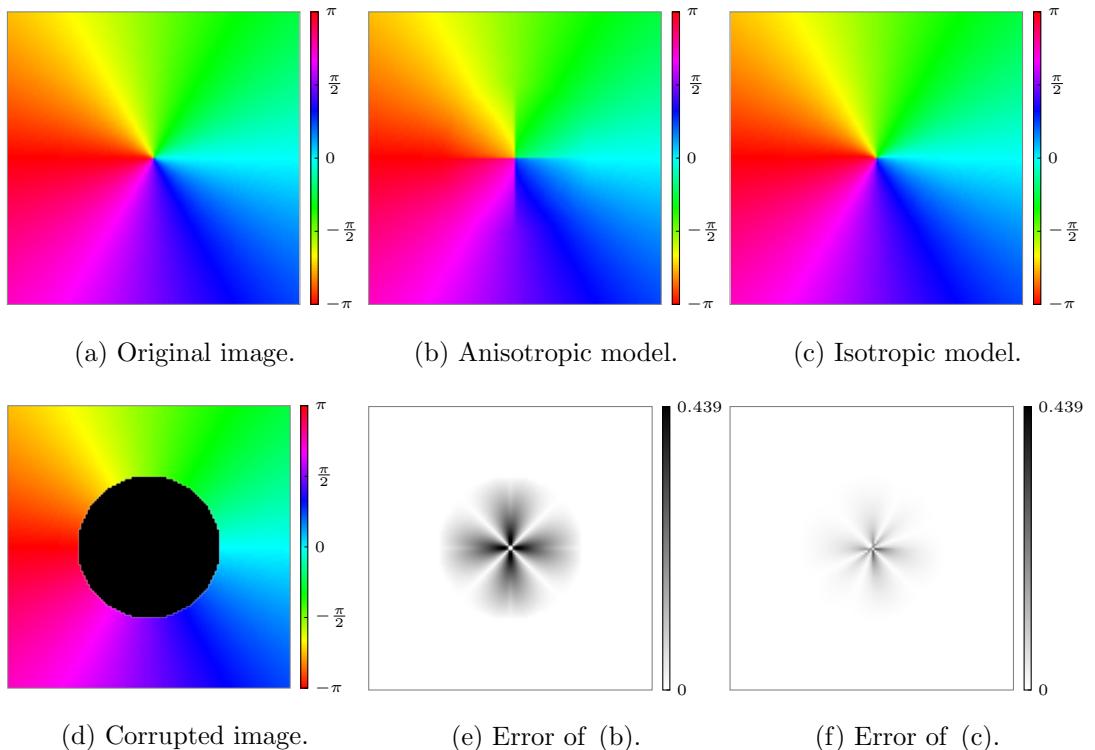
Next we want to demonstrate the difference between the anisotropic (12) and isotropic (13) half-quadratic minimization methods. To this end, the function  $\text{atan2}(x, y)$  was sampled over a regular grid  $[-\frac{1}{2}, \frac{1}{2}]^2$  with grid size  $\frac{1}{128}$ , resulting in Fig. 2 (a). Then we corrupt the image by removing a circular region from the center as shown in Fig. 2 (d). Using the anisotropic functional leads to Fig. 2 (b), where we observe artifacts in vertical and horizontal directions. The image produced by applying the isotropic functional in Fig. 2 (c) does not have this problem. This effect is also illustrated by the error plots in Figs. 2 (e) and 2 (f).

#### 4.2 $\mathbb{S}^2$ -valued data

In our first example we denoise color images in the chromaticity and brightness space. For an RGB image the brightness is given by the real positive numbers  $b := (R^2 + G^2 + B^2)^{\frac{1}{2}}$  and the chromaticity by the  $\mathbb{S}^2$ -values  $c := (R, G, B)/b$ . We want to mention, that the first two examples consider problems, which have values on the positive octant, only in the last



**Figure 1.** Restoration of a noisy cyclic signal by half-quadratic minimization with various functions  $\varphi$ . (a) original (red) and noisy (blue) signals. Restored signal (blue) using (b)  $\varphi_1$ ,  $\varepsilon = 6 \times 10^{-1}$ ,  $\lambda = 3.4$ , err = 0.1007, (c)  $\varphi_2$ ,  $\varepsilon = 5 \times 10^{-1}$ ,  $\lambda = 5.2$ , err = 0.1007, (d)  $\varphi_3$ ,  $\varepsilon = \frac{1}{\sqrt{2}}$ ,  $\lambda = 10$ , err = 0.1001, (e)  $\varphi_2$ ,  $\varepsilon = 1$ ,  $\lambda = 20$ , err = 0.1733, (f)  $\varphi_3$ ,  $\varepsilon = \sqrt{5}$ ,  $\lambda = 10$ , err = 0.3756.



**Figure 2.** Inpainting of an image with cyclic data using the anisotropic and the isotropic model. (a) original image, (d) corrupted image. Restoration with (b) the anisotropic model and (c) the isotropic model using the function  $\varphi_1$  with  $\lambda = 0.001$  and  $\varepsilon = 10^{-2}$ . (e) and (f) error images.



**Figure 3.** Denoising in the chromaticity-brightness space. (a) Original image “Peppers”. (b) Corrupted image by Gaussian noise on RGB,  $\sigma = 0.1$ . Restored images by (c)  $\varphi_1$  using  $\varepsilon_c = 10^{-3}$ ,  $\varepsilon_b = 10^{-2}$ ,  $\lambda_c = 0.44$ ,  $\lambda_b = 0.08$ , (d)  $\varphi_2$  using  $\varepsilon_c = 10^{-3}$ ,  $\varepsilon_b = 10^{-3}$ ,  $\lambda_c = 15$ ,  $\lambda_b = 10$ , (e)  $\varphi_3$  using  $\varepsilon_c = 2\sqrt{3}$ ,  $\varepsilon_b = \sqrt{23}$ ,  $\lambda_c = 0.1$ ,  $\lambda_b = 0.03$ , and (f) TV method in [56],  $\alpha = 0.05$ .

example we look at data covering the whole sphere. We compare half-quadratic minimization with the different functions  $\varphi$  and the TV approach from [56] in Fig. 3. We took the image “Peppers”<sup>1</sup> in Fig. 3 (a) and added Gaussian noise with standard deviation  $\sigma = 0.1$  to all three color channels in the RGB model. For denoising in the chromaticity-brightness model we optimized  $\lambda$  with respect to best PSNR for both channels separately using a grid search on  $\frac{1}{100}\mathbb{N}$  for  $\varphi_1$  and  $\varphi_3$ , and for  $\varphi_2$  on  $\mathbb{N}$ . Furthermore, we optimized  $\varepsilon$  first in order of magnitude  $k = 10^j$  and refined the search on  $\frac{k}{10}\mathbb{N}$ . Both channels are restored using the same function. The square-root functional  $\varphi_1$  shows smoother transitions at edges while the Huber function  $\varphi_2$  tends more to staircasing. Using the exponential function  $\varphi_3$  yields a worse PSNR and hence does not compete with the previous two functions; the edges are preserved but within the more constant regions some noise is left. The bright spot (see upper magnification) appears also too smooth. This originates from the too smooth transitions in the brightness which are not detected as edges. The TV regularization introduces staircasing and it is not able to reduce the noise in the dark area (lower magnification).

In the second example we use half-quadratic minimization for colorization in the chromaticity-brightness space. We assume that the brightness of the image is known, but 99 percent of the chromaticity information is lost. The original image is shown in Fig. 4 (a) and its corrupted version in Fig. 4 (b). For inpainting the chromaticity we have used a nearest neighbor initialization. With the regularizing function  $\varphi_1$  we obtain the result depicted in Fig. 4 (c). We compare this with Fig. 4 (d) which is obtained by using the chromaticity colorization method in [37] which we have implemented for comparison.

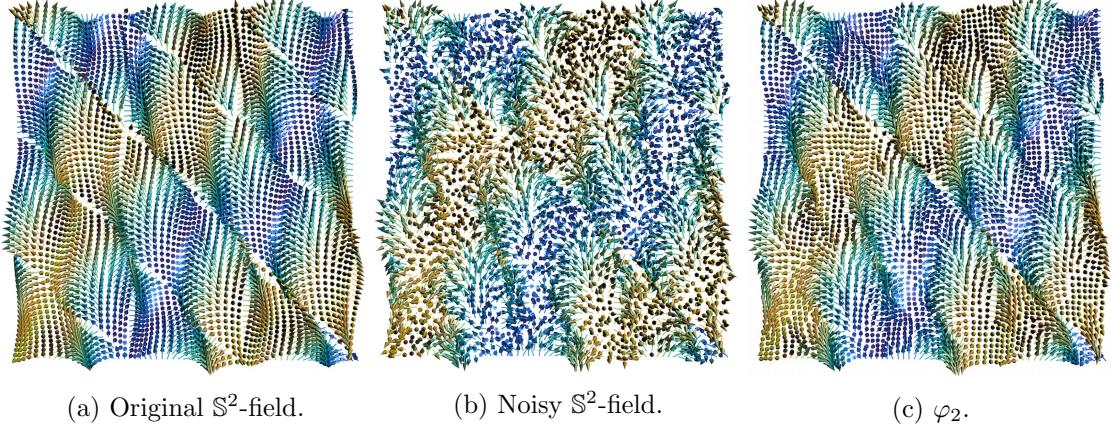
Our final experiment shows the smoothing of 3D directions in the synthetic image in

---

<sup>1</sup>Taken from the USC-SIPI Image Database, available online at <http://sipi.usc.edu/database/database.php?volume=misc&image=15>



**Figure 4.** Image colorization. (a) Original image. (b) Corrupted image where 99% of the color information (chromaticity) is lost. Colorization using (c) inpainting of the chromaticity with  $\varphi_1$ ,  $\lambda = 1$ ,  $\varepsilon = 10^{-1}$ . (d) the method in [37] with parameters  $r = 1, p = 1, \sigma_1 = 2, \sigma_2 = \infty, \gamma = 0$ .



**Figure 5.** (a) Original  $\mathbb{S}^2$ -field of size  $64 \times 64$ . (b) Corrupted field by Gaussian noise,  $\sigma = 0.1$ . (c) Restored field with  $\varphi_2$ ,  $\lambda = 2.6$ ,  $\varepsilon = 10^{-1}$ , leaving an error  $\text{err} = 0.1705$ .

Fig. 5 (b). We use half-quadratic minimization with  $\varphi_2$  to obtain Fig. 5 (c). The original pattern is again visible.

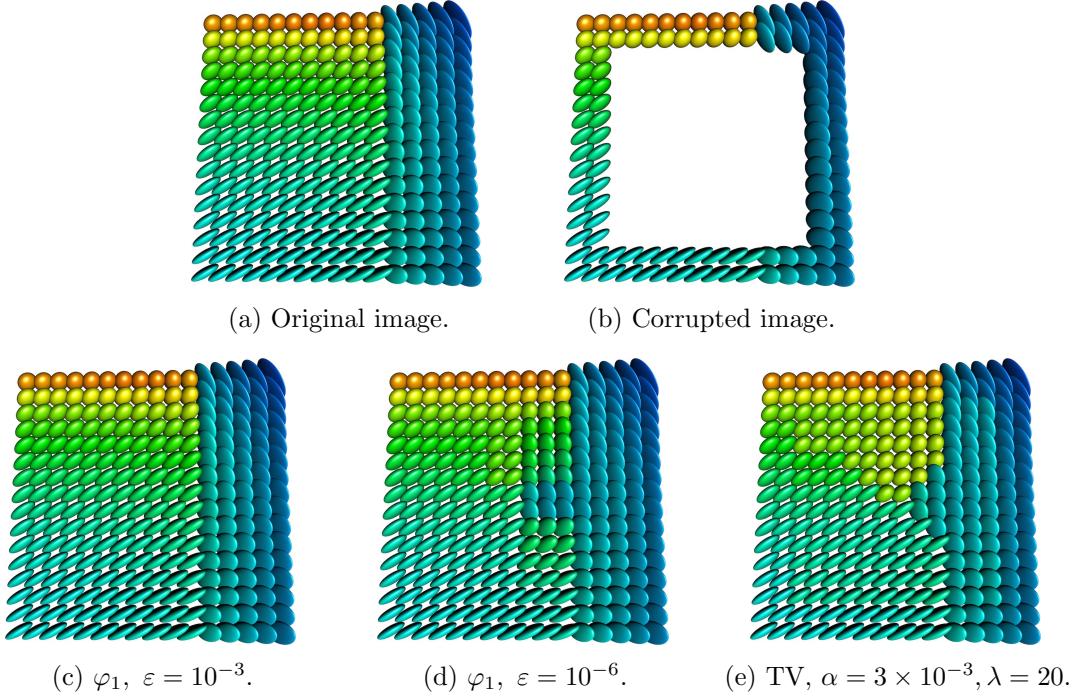
### 4.3 $\mathcal{P}(3)$ -valued data

Our first example illustrates the inpainting capabilities of the half-quadratic minimization method by an artificial example. The  $\mathcal{P}(3)$ -valued image of size  $16 \times 16$  in Fig. 6 (a) has a jump at  $\frac{2}{3}$  in  $x$  direction. We destroy a center square of size  $12 \times 12$ , see Fig. 6 (b). We are able to reconstruct the inpainting area nearly perfectly by using  $\varphi_1$  and  $\varepsilon = 10^{-3}$ , see Fig. 6 (c). Nevertheless decreasing  $\varepsilon$  introduces more and more staircasing, cf. Fig. 6 (d). This resembles the TV case shown in Fig. 6 (e), using the model from [56].

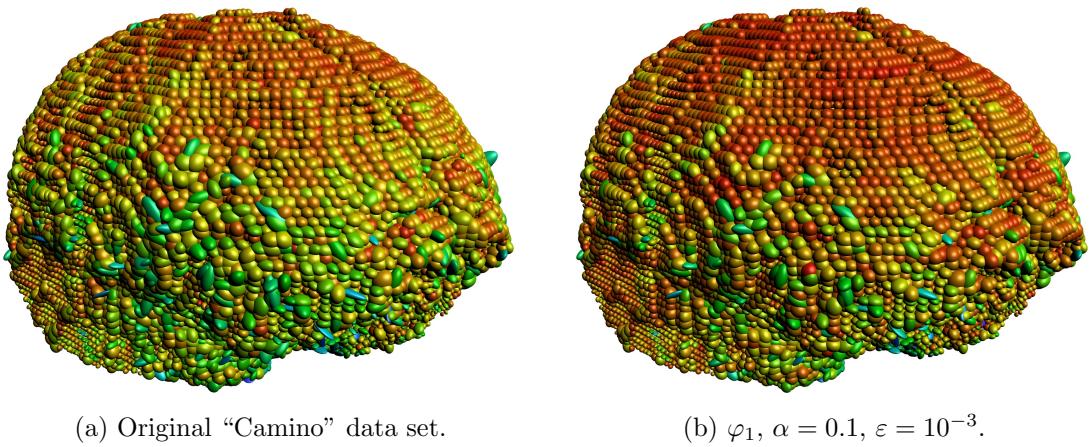
An important application of  $\mathcal{P}(3)$ -valued image denoising is Diffusion Tensor Magnetic Resonance Imaging (DT-MRI). The Camino project<sup>2</sup> [17] provides a DT-MRI dataset of the human head which is freely available.<sup>3</sup> The complete data is given as a 3D image  $\tilde{f} = (\tilde{f}_{i,j,k}) \in \mathcal{P}(3)^{112 \times 112 \times 50}$ , where we apply the half-quadratic minimization to each of the traversal planes  $k \in \{1, \dots, 50\}$ . The original dataset, cf. Fig. 7 (a), is plotted

<sup>2</sup>see <http://cmic.cs.ucl.ac.uk/camino>

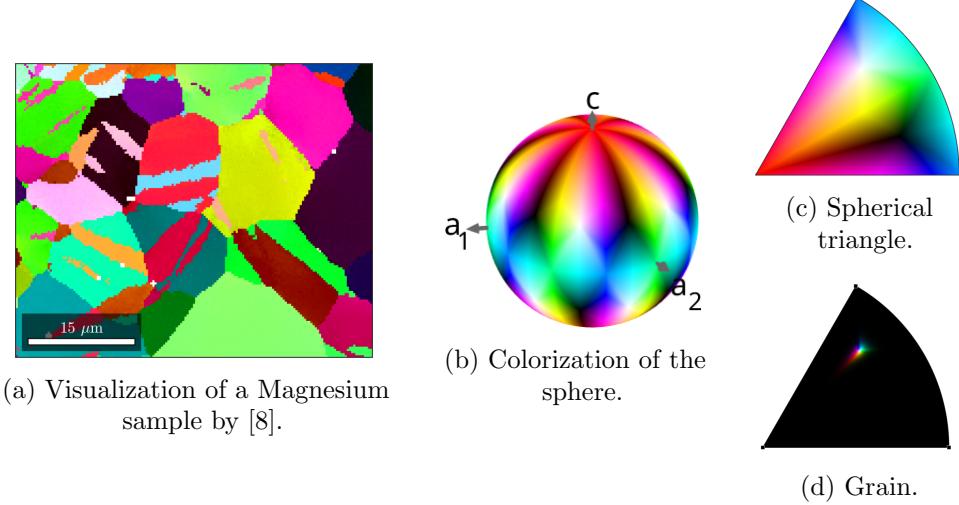
<sup>3</sup>follow the tutorial at <http://cmic.cs.ucl.ac.uk/camino//index.php?n=Tutorials.DTI>



**Figure 6.** Inpainting of an  $\mathcal{P}(3)$ -valued image. (a) Original image. (b) Image with unknown areas. Half-quadratic based inpainting (c) yields a perfect result. Decreasing  $\varepsilon$  like in (d) yields a result closer to (e) TV.



**Figure 7.** Denoising with half-quadratic minimization of every traversal plane of the ‘‘Camino’’ data set (DT-MRI of the human head).



**Figure 8.** (a) The raw EBSD data of a Magnesium sample, (b) the colorization of the sphere used to assign to each rotation a certain color according to the mapping  $\text{SO}(3)/S \ni m \mapsto m^{-1}(0, 0, 1)^T \in \mathbb{S}^2/S$ , (c) colorization of a spherical triangle, (d) stretched colorization for one grain.

using the anisotropy index relative to the Riemannian distance [32] normalized onto  $[0, 1)$  and colored in hue. The half-quadratic minimization is used with  $\varphi_1$  and the parameters  $\lambda = 0.1$ ,  $\varepsilon = 10^{-3}$  and a maximum change between two successive iterations being  $10^{-12}$  as a stopping criterion. We obtain the result shown in Fig. 7(b). For the complete dataset of 168,169 nonzero matrices, the algorithm needed 2,492 seconds to compute the result.

#### 4.4 SO(3)-valued data

Processing images with  $\text{SO}(3)$ -valued entries is fundamental in the analysis of polycrystalline materials by means of Electron Backscattered Diffraction (EBSD), cf. [2, 28]. Since the microscopic grain structure affects macroscopic attributes of materials such as ductility, electrical and lifetime properties, there is a growing interest in the grain structure of crystalline materials such as metals and minerals. EBSD provides us for each position on the surface of a specimen with a so called Kikuchi pattern, which allows the identification of the structure (material index) and the orientation of the crystal at this position relative to a fixed coordinate system ( $\text{SO}(3)$  value). Since the atomic structure of a crystal is invariant under its specific symmetry group  $S \subset \text{SO}(3)$  the orientation is only given as an equivalence class  $[m_0] = \{m_0 s \mid s \in S\} \in \text{SO}(3)/S$ ,  $m_0 \in \text{SO}(3)$ .

Fig. 8(a) displays a typical EBSD image consisting of lattice orientations of deformed Magnesium collected by [43]. Each pixel of the image corresponds to a position on the surface of a Magnesium specimen. The color of the pixels is chosen corresponding to the orientation measured at this position according to the following color mapping: for a fixed vector  $\vec{r} \in \mathbb{S}^2$  we consider the mapping  $\Phi: \text{SO}(3)/S \rightarrow \mathbb{S}^2/S$ ,  $[m] \mapsto [m^{-1}\vec{r}]$ . Next we colorize the quotient  $\mathbb{S}^2/S$  as it is depicted in Fig. 8(b). From the colorization scheme the symmetry group  $S \subset \text{SO}(3)$  of Magnesium becomes visible which has six rotations with respect to a 6-folded axis ( $k\pi/3$ ,  $k = 1, \dots, 6$  rotations around  $c$  direction) and six rotations with respect to 2-folded axis  $a_1$ ,  $a_2$  perpendicular to that.

EBSD images usually consist of regions with similar orientations called grains. For certain macroscopic properties the pattern of orientations within single grains is of importance [7], e.g., for the computation of geometrically necessary dislocations [35, 49] the

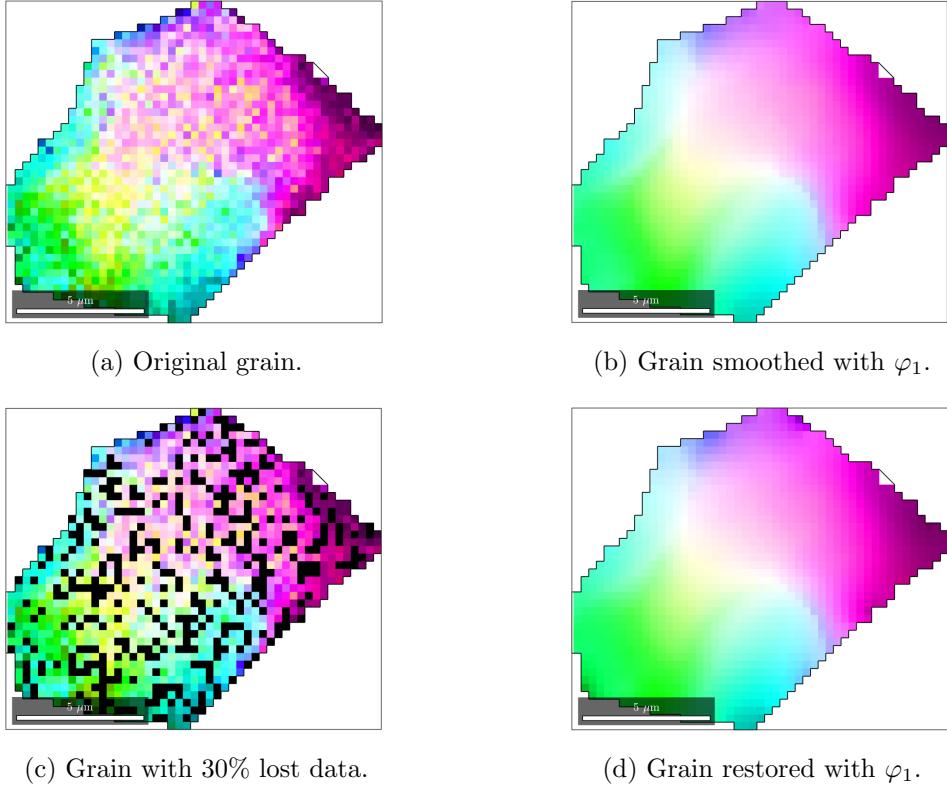
gradient of the rotations within single grains has to be determined. As the rotation determination by Kikuchi patterns is sometimes fragile, the rotation valued images determined by EBSD are often corrupted by noise and suffer from missing data so that denoising and inpainting techniques have to be applied [25]. For detecting grains in the raw EBSD data we applied a thresholding algorithm [8]. Fig. 9 (a) displays a single grain with its rotations. Since the rotations vary very little within a single grain we applied a sharper colorization, cf. Fig. 8 (d) to make the noise and the rotation gradient visible.

We want to apply half-quadratic minimization to denoise EBSD images. Since crystallographic symmetry groups are finite the quotient  $\text{SO}(3)/S$  is locally isomorphic to  $\text{SO}(3)$ . In particular, the formulas given in Appendix B.3 can be applied. In a first experiment we apply half-quadratic minimization using  $\varphi_1$  to the rotation-valued image depicted in Fig. 9 (a) which leads to the smooth image in Fig. 9 (b). In a second experiment we randomly removed 30% of the data shown in Fig. 9 (c). Using half-quadratic minimization for jointly inpainting and denoising the image we obtain the result shown in Fig. 9 (d) which looks very similar to those in Fig. 9 (b). In a third experiment we applied half-quadratic minimization simultaneously to several grains. The challenge from the mathematical point of view is that  $\text{SO}(3)$  is not an Hadamard manifold. Convergence can be guaranteed only locally, which is the case for single grains but may be not true when considering several grains simultaneously. From the practical point of view this case is especially interesting as missing data usually occur at grain boundaries, i.e., between grains. However, for our data we have got promising results. Fig. 10 (a) shows the grain from the previous example (pink color) and two other grains in its neighborhood. Note that the top middle area (light green) and top right area (brown-green) belong to the same grain. Pixels with missing data are plotted white. Half-quadratic minimization restoration with  $\varphi_1$  improves the image as can be seen in Fig. 10 (b). For our last experiment we again randomly remove 30% of the data, cf. Fig 10 (c). Restoration with  $\varphi_1$  leads to the result in Fig. 10 (d), which is again hardly to distinguish from Fig. 10 (b). Using  $\varphi_3$  leads to even better results as depicted in Fig. 10 (e). We can adjust the smoothing in such a way that the edge distinguishing two grains is not smoothed, while smaller rotation changes are smoothed. This is advantageous in the large top grain. Finally Fig. 10 (f) shows the zoom to the (pink) grain from Fig. 9 with adapted color map.

## 5 Conclusions

We adapted the principles of half-quadratic minimization to the setting of complete, connected Riemannian manifolds. In particular, the notation of the  $c$ -transform provides an interesting point of view. For Hadamard manifolds we proved the existence and uniqueness of the minimizer of the corresponding functionals as well as the convergence for the alternating minimization algorithm under moderate assumptions. The multiplicative half-quadratic minimization method resembles a quasi-Newton method [33] and appears to be very efficient in our numerical examples. There are numerous applications of the approach. In this paper, images having values in a manifold such as the 2-sphere or the symmetric positive definite matrices were denoised. The method was also used for inpainting missing information into images consisting of either rotation matrices or symmetric positive definite matrices. In the chromaticity-brightness color model, the inpainting technique was applied to the task of colorization. The method has further potential in EBSD.

Topics of future research are the derivation of convergence proofs for more general manifolds under special assumptions on the local behavior of the data. Furthermore, different data terms must be included for other applications, and the inclusion of higher



**Figure 9.** (a) Original grain from a Magnesium specimen (grain boundary in black), (b) smoothed grain using  $\varphi_1$  with parameters  $\lambda = 0.1$ ,  $\varepsilon = 10^{-2}$ , (c) grain with 30% lost data, marked in white, (d) inpainted and smoothed grain with  $\varphi_1$  and the same parameters.

order differences into the regularization term of the model is of interest.

## A Proofs

*Proof of Proposition 2.1.* 1. In the additive case we have

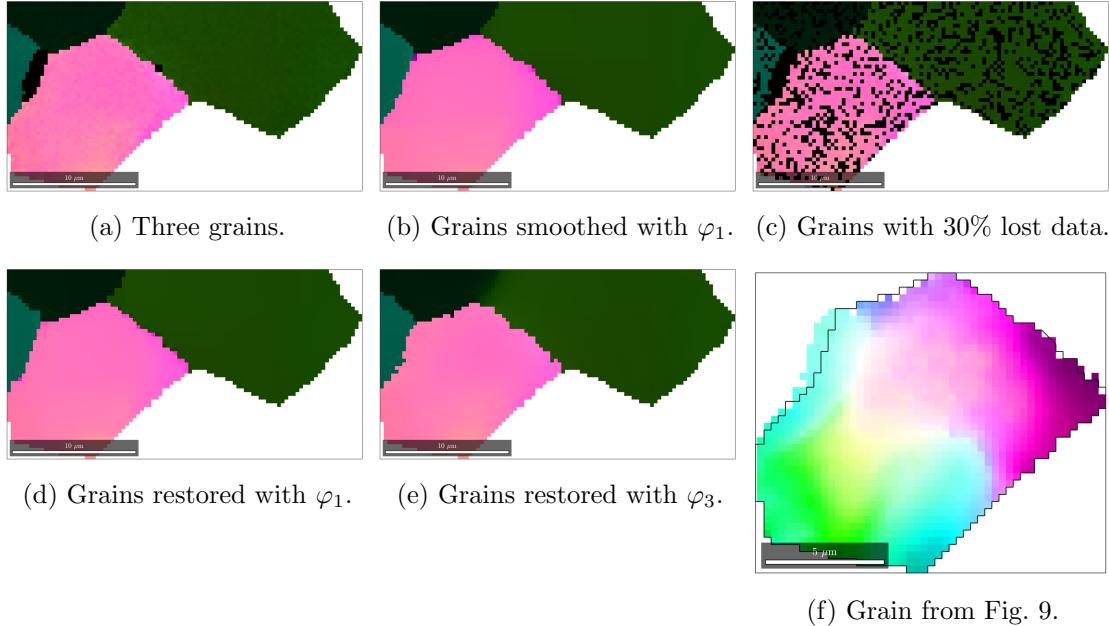
$$\begin{aligned} \varphi^c(s) &:= \inf_{t \in \mathbb{R}} \left\{ \frac{1}{2} \left( \sqrt{a}t - \frac{1}{\sqrt{a}}s \right)^2 - \varphi(t) \right\} = \inf_{t \in \mathbb{R}} \left\{ -st + \frac{1}{2}at^2 - \varphi(t) \right\} + \frac{1}{2a}s^2 \\ &= -\sup_{t \in \mathbb{R}} \left\{ st - \underbrace{\left( \frac{1}{2}at^2 - \varphi(t) \right)}_{\Phi} \right\} + \frac{1}{2a}s^2 \\ &= -\Phi^*(s) + \frac{1}{2a}s^2 \end{aligned}$$

By assumption on  $\Phi$  we know that  $\Phi = \Phi^{**}$  which implies

$$\begin{aligned} \varphi^{cc}(t) &= \inf_{s \in \mathbb{R}} \left\{ -st + \frac{1}{2a}s^2 - \varphi^c(s) \right\} + \frac{1}{2}at^2 = \inf_{s \in \mathbb{R}} \left\{ -ts + \Phi^*(s) \right\} + \frac{1}{2}at^2 \\ &= -\Phi(t) + \frac{1}{2}at^2 = \varphi(t). \end{aligned}$$

This finishes the proof of i). The function

$$h(t) := c(t, s) - \varphi(t) = \frac{1}{2}at^2 - \varphi(t) - st + \frac{1}{2a}s^2$$



**Figure 10.** (a) Three grains from a Magnesium specimen with missing data in white. (b) Smoothed (and ipainted) grains with  $\varphi_1$  and parameters  $\lambda = 0.15$ ,  $\varepsilon = 10^{-4}$ . (c) Grains with 30% data lost. Restored grains from (c) with (d)  $\varphi_1$  and parameters  $\lambda = 0.05$ ,  $\varepsilon = 0.5 \times 10^{-4}$ , (e)  $\varphi_3$  and parameters  $\lambda = 0.1$ ,  $\varepsilon = \sqrt{30}$ . (f) Grain from Fig. 9 after restoration with  $\varphi_3$  (original boundary in black).

is continuous, convex and by (10) coercive so that the global minimizer in (7) is attained for  $0 = h'(t) = at - \varphi'(t) - s$ , i.e., for  $(t, at - \varphi'(t))$  which proves ii).

2. In the multiplicative case we obtain, since  $\varphi$  is even,

$$\begin{aligned} \varphi^c(s) &:= \inf_{t \in \mathbb{R}} \{t^2 s - \varphi(t)\} = \inf_{t \geq 0} \{t^2 s - \varphi(t)\} \\ &= \inf_{r \geq 0} \{rs - \varphi(\sqrt{r})\} = -\sup_{r \geq 0} \{-rs - (-\varphi(\sqrt{r}))\} \\ &= -\Phi^*(-s). \end{aligned}$$

We have

$$\varphi^{cc}(t) = \varphi^{cc}(-t) = \inf_{s \in \mathbb{R}} \{t^2 s - \varphi^c(s)\}.$$

so that we can restrict our attention to  $t \geq 0$ . By assumption,  $\Phi$  is convex and lsc. Thus,  $\Phi = \Phi^{**}$  and we obtain for  $t \geq 0$  that

$$\varphi^{cc}(\sqrt{t}) = \inf_{s \in \mathbb{R}} \{ts - \varphi^c(s)\} = -\sup_{s \in \mathbb{R}} \{-ts - \Phi^*(-s)\} = -\Phi(t) = \varphi(\sqrt{t}).$$

This yields i). To see ii) we first note that condition (9) implies for  $s \geq 0$  that the objective function in (7) is coercive such that the infimum is attained. For  $s < 0$ , we have  $\varphi^c(s) = -\infty$ . For  $s \geq 0$ , we obtain

$$\arg \min_{t \geq 0} \{t^2 s - \varphi(t)\} = \left( \arg \min_{r \geq 0} \{rs - \varphi(\sqrt{r})\} \right)^{\frac{1}{2}}$$

By assumption on  $\Phi$ , the function  $h(r) := rs - \varphi(\sqrt{r})$  is convex in  $\mathbb{R}_{\geq 0}$ . A global minimizer of  $h$  is attained either for the solution of  $0 = h'(r) = s - \frac{1}{2\sqrt{r}}\varphi'(\sqrt{r})$  if this solution is positive

or for  $r = 0$  if  $\lim_{r \rightarrow 0^+} h'(r) \geq 0$ , i.e.,  $s \geq \frac{1}{2}\varphi''(0+)$ . Therefore  $(t, s) = (0, \frac{1}{2}\varphi''(0+))$  is a solution.

Finally, the concavity of  $\varphi(\sqrt{t})$  for  $t \geq 0$  implies that  $\varphi'(\sqrt{t})/(2\sqrt{t})$  and thus  $\varphi(t)/(2t)$  is decreasing. Under the additional assumption in iii) we get  $s \in (0, \frac{\varphi''(0+)}{2}]$ .  $\square$

*Proof of Proposition 3.3.* i) If  $\mathcal{V} = \mathcal{G}$ , let  $\|d(u, f)\|_2 \rightarrow \infty$ , where  $d(u, f) := (d(u_i, f_i))_{i \in \mathcal{G}}$ . Then  $F(u) := \frac{1}{2} \sum_{i \in \mathcal{V}} d^2(u_i, f_i)$  goes to infinity and the functionals  $J_\nu$ ,  $\nu = 1, 2$ , are coercive. In the case  $\mathcal{V} \neq \mathcal{G}$  choose  $i_0 \in \mathcal{V}$  and let  $u_0$  be the constant image with entries  $f_{i_0}$ . Let  $\|d(u, u_0)\|_2 \rightarrow \infty$ . Assume that  $J_\nu(u)$  remains finite, so that in particular  $d(f_{i_0}, u_{i_0})$  and  $d(u_i, u_j)$ ,  $j \in \mathcal{N}(i)^+$ ,  $i \in \mathcal{G}$  are finite. By the construction of the neighborhoods  $\mathcal{N}(i)^+$  there exists for every  $j \in \mathcal{G}$  a path  $i_0, i_1, \dots, i_{k_j} = j$  with  $i_{l+1} \in \mathcal{N}(i_l)^+$  and

$$d(f_{i_0}, u_j) \leq d(f_{i_0}, u_{i_0}) + d(u_{i_0}, u_{i_1}) + \dots + d(u_{i_{k_j-1}}, u_j).$$

Since the right-hand side remains finite this contradicts  $\|d(u, u_0)\|_2 \rightarrow \infty$ . Hence  $J_\nu$ ,  $\nu = 1, 2$  are coercive.

ii) By (D2) we have that  $F(u)$  is convex and strictly convex if  $\mathcal{V} = \mathcal{G}$ . If a function  $h: \mathcal{H}^\kappa \rightarrow \mathbb{R}$  is convex, then, for any geodesic  $\gamma: [0, 1] \rightarrow \mathcal{H}^\kappa$  joining  $x, y \in \mathcal{H}^\kappa$ , we obtain since  $\varphi$  is increasing and convex that

$$\varphi \circ h(\gamma(t)) \leq \varphi(th(x) + (1-t)h(y)) \leq t(\varphi \circ h)(x) + (1-t)(\varphi \circ h)(y) \quad (24)$$

so that  $\varphi \circ h$  is convex. If  $\varphi$  is strictly convex the last inequality is strong so that  $\varphi \circ h$  is strictly convex. With  $h := h_{ij} = d(u_i, u_j) : \mathcal{H}^2 \rightarrow \mathbb{R}$  this implies by (D1) that  $J_1$  is convex, resp. strictly convex. Concerning  $J_2$  notice that the convexity of  $h_i(x_0, x_i)$ ,  $i = 1, \dots, \kappa - 1$ , on  $\mathcal{H}^2$  implies convexity of  $h(x_0, \dots, h_{\kappa-1}) := (\sum_{i=1}^{\kappa-1} h_i^2(x_0, x_i))^{\frac{1}{2}}$  on  $\mathcal{H}^\kappa$  by

$$\begin{aligned} h^2(\gamma(t)) &= \sum_{i=1}^{\kappa-1} h_i^2(\gamma_0(t), \gamma_i(t)) \leq \sum_{i=1}^{\kappa-1} \left( th_i(\gamma_0(0), \gamma_i(0)) + (1-t)h_i(\gamma_0(1), \gamma_i(1)) \right)^2 \\ &= t^2 h^2(\gamma_0(0)) + (1-t)^2 h^2(\gamma_0(1)) + 2t(t-1) \sum_{i=1}^{\kappa-1} h_i(\gamma_0(0), \gamma_i(0)) h_i(\gamma_0(1), \gamma_i(1)) \end{aligned}$$

and by the Schwarz inequality

$$\begin{aligned} h^2(\gamma(t)) &\leq t^2 h^2(\gamma_0(0)) + (1-t)^2 h^2(\gamma_0(1)) + 2t(t-1)h(\gamma_0(0))h(\gamma_0(1)) \\ &= \left( th(\gamma_0(0)) + (1-t)h(\gamma_0(1)) \right)^2. \end{aligned}$$

In the case  $\mathcal{V} = \mathcal{G}$ , strict convexity follows by the strict convexity of the data term  $\sum_{i \in \mathcal{G}} d^2(f_i, u_i)$  and for strictly convex  $\varphi$  by the strict convexity in (24).  $\square$

*Proof of Theorem 3.4.* By Remark 3.1 we know that  $\lim_{k \rightarrow \infty} \mathcal{J}_\nu(u^{(k)}, v^{(k)}) =: \bar{b}$  and that there exists a subsequence  $\{(u^{(k_j)}, v^{(k_j)})\}_j$  which converges to some  $(\bar{u}, \bar{v})$ . Since  $\mathcal{J}_\nu$  is continuous we have  $\lim_{j \rightarrow \infty} \mathcal{J}_\nu(u^{(k_j)}, v^{(k_j)}) = \mathcal{J}_\nu(\bar{u}, \bar{v}) = \bar{b}$ . Let

$$\tilde{v} := s(d_{\bar{u}}) = \arg \min_v \mathcal{J}_\nu(\bar{u}, v) \quad \text{and} \quad \tilde{u} := \arg \min_u \mathcal{J}_\nu(u, \tilde{v}).$$

The continuity of  $s$  and  $d$  implies that  $\lim_{j \rightarrow \infty} v^{(k_j+1)} = \lim_{j \rightarrow \infty} s(d_{u^{(k_j)}}) = s(d_{\bar{u}}) = \tilde{v}$  and the continuity of  $T(u) := \arg \min_z \mathcal{J}_\nu(z, s(d_u))$  that  $\lim_{j \rightarrow \infty} u^{(k_j+1)} = \lim_{j \rightarrow \infty} T(u^{(k_j)}) = T(\bar{u}) = \tilde{u}$ . By (22) we conclude

$$\bar{b} = \lim_{j \rightarrow \infty} \mathcal{J}_\nu(u^{(k_j+1)}, v^{(k_j+1)}) = \mathcal{J}_\nu(\tilde{u}, \tilde{v}) \leq \mathcal{J}_\nu(\bar{u}, \tilde{v}) \leq \mathcal{J}_\nu(\bar{u}, \bar{v}) = \bar{b}.$$

Thus,  $\mathcal{J}_\nu(\tilde{u}, \tilde{v}) = \mathcal{J}_\nu(\bar{u}, \tilde{v}) = \mathcal{J}_\nu(\bar{u}, \bar{v})$  and since  $\mathcal{J}_\nu(\bar{u}, \cdot)$  and  $\mathcal{J}_\nu(\cdot, \tilde{v})$  have unique minimizers we obtain that  $\tilde{u} = \bar{u}$  and  $\tilde{v} = \bar{v}$ . Consequently,  $\bar{v} = s(d_{\bar{u}})$  and  $\bar{u} = \tilde{u} = \arg \min_u \mathcal{J}_\nu(u, s(d_{\bar{u}}))$ , which by Remark 2.2 implies that  $\bar{u}$  is a minimizer of  $J_\nu$ .

Assume that  $\lim_{k \rightarrow \infty} u^{(k)} \not\rightarrow \bar{u}$ . Then there exists  $\varepsilon > 0$  such that infinitely many  $u^{(k_i)}$  not contained in the open ball  $B_\varepsilon(\bar{u})$  of radius  $\varepsilon$  centered at  $\bar{u}$ . Since  $\{u^{(k_i)}\}$  is bounded, there exists a convergent subsequence  $u^{(k_{i_j})}$  which converges to some point  $u^* \neq \bar{u}$  in the closed set  $M \setminus B_\varepsilon$ . Then  $\lim_{j \rightarrow \infty} J_\nu(u^{(k_{i_j})}) = J_\nu(u^*) = J(\bar{u})$  which contradicts the fact that  $J_\nu$  has a unique minimizer.  $\square$

## B Exponential and Logarithmic Maps

### B.1 The Sphere $\mathbb{S}^2$

We use the parametrization

$$x(\theta, \varphi) := \begin{pmatrix} \cos \varphi \cos \theta \\ \sin \varphi \cos \theta \\ \sin \theta \end{pmatrix}, \quad \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \varphi \in [0, 2\pi).$$

Then we have the tangent spaces

$$T_x(\mathbb{S}^2) = T_{x(\theta, \varphi)}(\mathbb{S}^2) := \{\eta \in \mathbb{R}^{d+1} : \eta^\top x = 0\} = \text{span}\{e_1(\theta, \varphi), e_2(\theta, \varphi)\}$$

with the normed orthogonal vectors

$$e_1(\theta, \varphi) := \frac{\partial x}{\partial \theta} = \begin{pmatrix} -\cos \varphi \sin \theta \\ -\sin \varphi \sin \theta \\ \cos \theta \end{pmatrix}, \quad e_2(\theta, \varphi) := \frac{1}{\cos \theta} \frac{\partial x}{\partial \varphi} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}.$$

The Riemannian metric is just the Euclidean distance in  $\mathbb{R}^3$ . The geodesic distance is given by  $d_{\mathbb{S}^2}(x_1, x_2) := \arccos \langle x_1, x_2 \rangle$ , and the exponential map and (locally) its inverse, resp., by

$$\begin{aligned} \exp_x(t\eta) &:= \cos(t\|\eta\|_2)x + \sin(t\|\eta\|_2) \frac{\eta}{\|\eta\|_2}, \\ \log_{x_1} x_2 &:= \frac{x_2 - \langle x_1, x_2 \rangle x_1}{\|x_2 - \langle x_1, x_2 \rangle x_1\|_2} \arccos \langle x_1, x_2 \rangle. \end{aligned}$$

### B.2 The Manifold $\mathcal{P}(r)$ of Symmetric Positive Definite Matrices

By Exp and Log we denote the matrix exponential and logarithm defined by

$$\text{Exp } x := \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \quad \text{Log } x := \sum_{k=1}^{\infty} \frac{1}{k} (I - x)^k, \quad \rho(I - x) < 1.$$

Let  $\text{Sym}(r)$  denote the space of symmetric  $r \times r$  matrices with (Frobenius) inner product and norm

$$\langle A, B \rangle := \sum_{i,j=1}^r a_{ij} b_{ij}, \quad \|A\| := \left( \sum_{i,j=1}^r a_{ij}^2 \right)^{\frac{1}{2}}. \quad (25)$$

Let  $\mathcal{P}(r)$  be the manifold of symmetric positive definite  $r \times r$  matrices. It has the dimension  $\dim \mathcal{P}(r) = n = \frac{r(r+1)}{2}$ . The tangent space of  $\mathcal{P}(r)$  at  $x \in \mathcal{P}(r)$  is given by  $T_x \mathcal{P}(r) =$

$\{x\} \times \text{Sym}(r) := \{x^{\frac{1}{2}}\eta x^{\frac{1}{2}} : \eta \in \text{Sym}(r)\}$ , in particular  $T_I \mathcal{P}(r) = \text{Sym}(r)$ , where  $I$  denotes the  $r \times r$  identity matrix. The Riemannian metric on  $T_x \mathcal{P}$  reads

$$\langle \eta_1, \eta_2 \rangle_x := \text{tr}(\eta_1 x^{-1} \eta_2 x^{-1}) = \langle x^{-\frac{1}{2}} \eta_1 x^{-\frac{1}{2}}, x^{-\frac{1}{2}} \eta_2 x^{-\frac{1}{2}} \rangle,$$

where  $\langle \cdot, \cdot \rangle$  denotes the matrix inner product (25). The geodesic distance is given by  $d_{\mathcal{P}}(x_1, x_2) := \|\text{Log}(x_1^{-\frac{1}{2}} x_2 x_1^{-\frac{1}{2}})\|$ , and the exponential map and its inverse by

$$\exp_x(t\eta) := x^{\frac{1}{2}} \text{Exp}(tx^{-\frac{1}{2}} \eta x^{-\frac{1}{2}}) x^{\frac{1}{2}}, \quad \text{resp.}, \quad \log_{x_1} x_2 := x_1^{\frac{1}{2}} \text{Log}(x_1^{-\frac{1}{2}} x_2 x_1^{-\frac{1}{2}}) x_1^{\frac{1}{2}},$$

see [44].

### B.3 The Manifold $\text{SO}(3)$ of Rotation Matrices in $\mathbb{R}^3$

The manifold of  $3 \times 3$  rotation matrices is defined as

$$\text{SO}(3) := \{x \in \mathbb{R}^{3,3} \mid x^T x = I \text{ and } \det x = 1\}.$$

The geodesic distance between two rotation matrices  $x_1, x_2 \in \text{SO}(3)$  is given by

$$d_{\text{SO}(3)}(x_1, x_2) := \sqrt{2} \arccos\left(\frac{1 - \text{tr}(x_1^T x_2)}{2}\right).$$

The tangential space at  $x \in \text{SO}(3)$  reads

$$T_x \text{SO}(3) := \{xv : v \in T_I \text{SO}(3)\}, \quad T_I \text{SO}(3) := \{v \in \mathbb{R}^{3,3} : v + v^T = 0\}.$$

For  $\eta \in T_x \text{SO}(3)$  the exponential map at  $x \in \text{SO}(3)$  and (locally) its inverse are defined as

$$\exp_x(\eta) := x \text{Exp}(x^T \eta), \quad \text{resp.}, \quad \log_{x_1} x_2 = x_1 \text{Log}(x_1^T x_2).$$

The  $\text{SO}(3)$  can be parametrized in various ways. Due to the form of our data we prefer to use quaternions for the representation and the similarity of  $\text{SO}(3)$  to the group  $\mathbb{S}^3$ , see, e.g., [23]: We decompose the unit quaternions  $q = (s, v^T)^T \in \mathbb{S}^3$  into a real part  $s \in \mathbb{R}$  and a vector part  $v \in \mathbb{R}^3$ . The multiplication of two quaternions  $q_1, q_2 \in \mathbb{S}^3$  is given by

$$q_1 \circ q_2 := \begin{pmatrix} s_1 \\ v_1 \end{pmatrix} \circ \begin{pmatrix} s_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} s_1 s_2 - v_1^T v_2 \\ s_2 v_1 + s_1 v_2 + v_1 \times v_2 \end{pmatrix},$$

with the conjugate quaternion  $\bar{q} := (s, -v^T)^T$  as inverse element and unit element  $(1, 0, 0, 0)^T$ . The quaternions can be identified with the rotations of  $\text{SO}(3)$ , where the quaternions  $q$  and  $-q$  correspond to the same rotation. More precisely,  $\mathbb{S}^3$  is a double cover of  $\text{SO}(3)$ , see, [12, Chap. III, Sect. 10]:

$$\mathbb{S}_*^3 := \mathbb{S}^3 / \{-1, 1\} \cong \text{SO}(3).$$

We work with the representative having a positive first component. With this representation at hand, the Riemannian metric of  $\text{SO}(3)$  can be deduced from the Euclidean metric in  $\mathbb{R}^4$ . The geodesic distance can be written as

$$d_{\mathbb{S}_*^3}(q_1, q_2) = 2 \arccos |\langle q_1, q_2 \rangle|.$$

The exponential map of  $\eta \in T_p \mathbb{S}^3_*$  is given by

$$\exp_q \eta := \operatorname{sgn} s \begin{pmatrix} s \\ v \end{pmatrix}, \quad (s, v)^T := q \cos \|\eta\|_2 + \frac{\eta}{\|\eta\|_2} \sin \|\eta\|_2$$

and the logarithmic map at  $q_1$  of  $q_2$  by

$$\log_{q_1} q_2 := \frac{q_2 - \langle q_2, q_1 \rangle q_1}{\|q_2 - \langle q_2, q_1 \rangle q_1\|_2} \arccos(|\langle q_1, q_2 \rangle|) \operatorname{sgn}(\langle q_1, q_2 \rangle).$$

**Acknowledgement:** Funding by the DFG within the RTG GrK 1932 “Stochastic Models for Innovations in the Engineering Sciences”, project area P3, and in the projects STE 571/13-1 and BE 5888/2-1 is gratefully acknowledged. Furthermore RC gratefully acknowledges support by the HKRGC Grants No. CUHK300614, CUHK2/CRF/11G, AoE/M-05/12; CUHK DAG No. 4053007, and FIS Grant No. 1907303.

## References

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton and Oxford, Princeton University Press, 2008.
- [2] B. L. Adams, S. I. Wright, and K. Kunze. Orientation imaging: The emergence of a new microscopy. *Journal Metallurgical and Materials Transactions A*, 24:819–831, 1993.
- [3] A. D. Aleksandrov. A theorem on triangles in a metric space and some of its applications. In *Trudy Mat. Inst. Steklov.*, v 38, pages 5–23. Izdat. Akad. Nauk SSSR, Moscow, 1951.
- [4] M. Allain, J. Idier, and Y. Goussard. On global and local convergence of half-quadratic algorithms. *IEEE Transactions on Image Processing*, 15(5):1130–1142, 2006.
- [5] M. Bačák. *Convex Analysis and Optimization in Hadamard Spaces*, volume 22 of *De Gruyter Series in Nonlinear Analysis and Applications*. De Gruyter, Berlin, 2014.
- [6] M. Bačák, R. Bergmann, G. Steidl, and A. Weinmann. A second order non-smooth variational model for restoring manifold-valued images. *Preprint Univ. Kaiserslautern*, 2015.
- [7] F. Bachmann, R. Hielscher, P. E. Jupp, W. Pantleon, H. Schaeben, and E. Wegert. Inferential statistics of electron backscatter diffraction data from within individual crystalline grains. *Journal of Applied Crystallography*, 43:1338–1355, 2010.
- [8] F. Bachmann, R. Hielscher, and H. Schaeben. Grain detection from 2d and 3d EBSD data – specification of the MTEX algorithm. *Ultramicroscopy*, 111:1720–1733, 2011.
- [9] R. Bergmann, F. Laus, G. Steidl, and A. Weinmann. Second order differences of cyclic data and applications in variational denoising. *SIAM Journal on Imaging Sciences*, 7(4):2916–2953, 2014.
- [10] R. Bergmann and A. Weinmann. Inpainting of cyclic data using first and second order differences. In X.-C. Tai, E. Bai, T. Chan, S. Y. Leung, and M. Lysaker, editors, *EMMCVPR2015*, Lecture Notes in Computer Science, pages 155–168, Berlin, 2015. Springer.

- [11] R. Bergmann and A. Weinmann. A second order TV-type approach for inpainting and denoising higher dimensional combined cyclic and vector space data. *ArXiv Preprint*, 1501.02684, 2015.
- [12] G. E. Bredon. *Topology and Geometry*, volume 139 of *Graduate Texts in Mathematics*. Springer, New York, 1993.
- [13] R. Bürgmann, P. A. Rosen, and E. J. Fielding. Synthetic aperture radar interferometry to measure earth's surface topography and its deformation. *Annual Reviews Earth and Planetary Science*, 28(1):169–209, 2000.
- [14] F. Champagnat and J. Idier. A connection between half-quadratic criteria and EM algorithms. *IEEE Signal Processing Letters*, 11(9):709–712, 2004.
- [15] T. F. Chan, S. Kang, and J. Shen. Total variation denoising and enhancement of color images based on the CB and HSV color models. *Journal of Visual Communication and Image Representation*, 12:422–435, 2001.
- [16] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2):298–311, 1997.
- [17] P. A. Cook, Y. Bai, S. Nedjati-Gilani, K. K. Seunarine, M. G. Hall, G. J. Parker, and D. C. Alexander. Camino: Open-source diffusion-mri reconstruction and processing. In *Proc. Intl. Soc. Mag. Reson. Med. 14*, page 2759, Seattle, WA, USA, 2006.
- [18] I. Daubechies, R. DeVore, and C. S. Güntürk. Iteratively reweighted least squares minimization for sparse recovery. *Communications in Pure and Applied Mathematics*, 63(1):1–38, 2010.
- [19] A. H. Delaney and Y. Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Transactions on Image Processing*, 7(2):204–221, 1998.
- [20] C.-A. Deledalle, L. Denis, and F. Tupin. NL-InSAR: Nonlocal interferogram estimation. *IEEE Transactions on Geoscience Remote Sensing*, 49(4):1441–1452, 2011.
- [21] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3):367–383, 1992.
- [22] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, 4(7):932–946, 1995.
- [23] M. Gräf. A unified approach to scattered data approximation on  $\mathbb{S}^3$  and  $\text{SO}(3)$ . *Advances in Computational Mathematics*, 37:379–392, 2012.
- [24] P. Grohs and M. Sprecher. Total variation regularization by iteratively reweighted least squares on Hadamard spaces and the sphere. *Preprint 2014-39, ETH Zürich*, 2014.
- [25] V. K. Gupta and S. R. Agnew. A simple algorithm to eliminate ambiguities in EBSD orientation map visualization and analyses: Application to fatigue crack-tips/wakes in aluminum alloys. *Microscopy and Microanalysis*, 16:831, 2010.

- [26] J. Jost. *Nonpositive Curvature: Geometric and Analytic Aspects*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1997.
- [27] R. Kimmel and N. Sochen. Orientation diffusion or how to comb a porcupine. *Journal of Visual Communication and Image Representation*, 13:238–248, 2002.
- [28] K. Kunze, S. I. Wright, B. L. Adams, and D. J. Dingley. Advances in automatic EBSP single orientation measurements. *Textures and Microstructures*, 20:41 – 54, 1993.
- [29] R. Lai and S. Osher. A splitting method for orthogonality constrained problems. *Journal of Scientific Computing*, 58(2):431–449, 2014.
- [30] C. L. Lawson. Contributions to the theory of linear least maximum approximation. *Ph.D. Thesis, University of California, Los Angeles*, 1961.
- [31] J. Lellmann, E. Strekalovskiy, S. Koetter, and D. Cremers. Total variation regularization for functions with values in a manifold. In *IEEE ICCV 2013*, pages 2944–2951, 2013.
- [32] M. Moakher and P. G. Batchelor. Symmetric positive-definite matrices: From geometry to applications and visualization. In J. Weickert and H. Hagen, editors, *Visualization and Processing of Tensor Fields*, pages 285–298. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
- [33] M. Nikolova and R. H. Chan. The equivalence of half-quadratic minimization and the gradient linearization iteration. *IEEE Transactions on Image Processing*, 16(6):1623–1627, 2007.
- [34] M. Nikolova and M. K. Ng. Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM Journal on Scientific Computing*, 27(3):937–966, 2005.
- [35] J. F. Nye. Some geometrical relations in dislocated crystals. *Acta Mater.*, 1:153–162, 1953.
- [36] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66:41–66, 2006.
- [37] M. H. Quang, S. H. Kang, and T. M. Le. Image and video colorization using vector-valued reproducing kernel Hilbert spaces. *Journal of Mathematical Imaging and Vision*, 37:49–65, 2010.
- [38] M. Raptis and S. Soatto. Tracklet descriptors for action modeling and video analysis. In *ECCV 2010*, pages 577–590. Springer, 2010.
- [39] J. G. Rešetnjak. Non-expansive maps in a space of curvature no greater than  $K$ . *Akademija Nauk SSSR. Sibirskoe Otdelenie. Sibirskiĭ Matematicheskiĭ Žurnal*, 9:918–927, 1968.
- [40] G. Rosman, M. Bronstein, A. Bronstein, A. Wolf, and R. Kimmel. Group-valued regularization framework for motion segmentation of dynamic non-rigid shapes. In *Scale Space and Variational Methods in Computer Vision*, pages 725–736. Springer, 2012.

- [41] G. Rosman, X.-C. Tai, R. Kimmel, and A. M. Bruckstein. Augmented-Lagrangian regularization of manifold-valued maps. *Methods and Applications of Analysis*, 21(1):105–122, 2014.
- [42] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1):259–268, 1992.
- [43] Z.-Z. Shi and J.-S. Lecomte. private communication, 2014.
- [44] S. Sra and R. Hosseini. Conic geometric optimization on the manifold of positive definite matrices. *ArXiv Preprint 1320.1039v3*, 2014.
- [45] G. Steidl, S. Setzer, B. Popilka, and B. Burgeth. Restoration of matrix fields by second order cone programming. *Computing*, 81:161–178, 2007.
- [46] E. Strekalovskiy and D. Cremers. Total variation for cyclic structures: Convex relaxation and efficient minimization. In *IEEE CVPR 2011*, pages 1905–1911. IEEE, 2011.
- [47] E. Strekalovskiy and D. Cremers. Total cyclic variation and generalizations. *Journal of Mathematical Imaging and Vision*, 47(3):258–277, 2013.
- [48] K. T. Sturm. Probability measures on metric spaces of nonpositive curvature, heat kernels and analysis on manifolds, graphs, and metric spaces. *Contemporary Mathematics*, 338:357–390, 2003.
- [49] S. Sun, B. Adams, and W. King. Observation of lattice curvature near the interface of a deformed aluminium bicrystal. *Phil. Mag. A*, 80:9–25, 2000.
- [50] O. Tuzel, F. Porikli, and P. Meer. Learning on Lie groups for invariant detection and tracking. In *CVPR 2008*, pages 1–8. IEEE, 2008.
- [51] L. Vese and S. Osher. Numerical methods for p-harmonic flows and applications to image processing. *SIAM Journal on Numerical Analysis*, 40:2085–2104, 2002.
- [52] C. Villani. *Topics in Optimal Transportation*. AMS, Providence, 2003.
- [53] C. R. Vogel and M. E. Oman. Iterative method for total variation denoising. *SIAM Journal on Scientific Computing*, 17(1):474–477, 1996.
- [54] C. R. Vogel and M. E. Oman. Fast, robust total variation-based reconstruction of noisy, blurred images. *IEEE Transactions on Image Processing*, 7(6):813–824, 1998.
- [55] J. Weickert, C. Feddern, M. Welk, B. Burgeth, and T. Brox. PDEs for tensor image processing. In J. Weickert and H. Hagen, editors, *Visualization and Processing of Tensor Fields*, pages 399–414, Berlin, 2006. Springer.
- [56] A. Weinmann, L. Demaret, and M. Storath. Total variation regularization for manifold-valued data. *SIAM Journal on Imaging Sciences*, 7(4):2226–2257, 2014.
- [57] M. Welk, C. Feddern, B. Burgeth, and J. Weickert. Median filtering of tensor-valued images. In B. Michaelis and G. Krell, editors, *Pattern Recognition*, Lecture Notes in Computer Science, 2781, pages 17–24, Berlin, 2003. Springer.