NTNU

Norwegian University of Science and Technology

The Riemannian Chambolle-Pock Algorithm

Ronny Bergmann

joint work with
R. Herzog, M. Silva Louzeiro, D. Tenbrinck, J. Vidal-Núñez.

Jahrestagung der Gesellschaft für angewandte Mathematik und Mechanik,

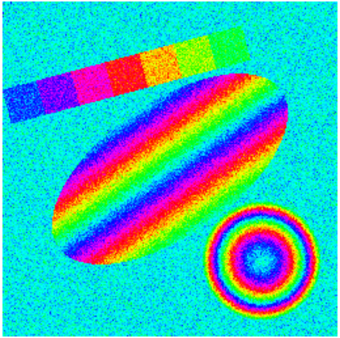
S21. Mathematical signal and image processing

August 17, 2022

Manifold-valued Signal & Image Processing

Tasks in image processing are often phrased as an optimisation problem.
Here. The pixel take values on a manifold

- phase-valued data ($\mathbb{S}^1$)
- wind-fields, GPS ($\mathbb{S}^2$)
- DT-MRI ($\mathcal{P}(3)$)
- EBSD, (grain) orientations ($\mathrm{SO}(n)$)



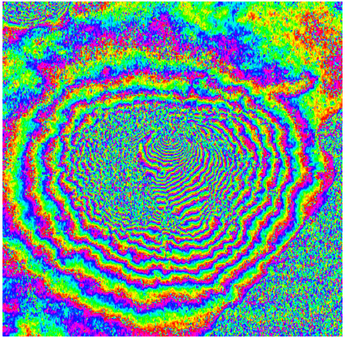
Artificial noisy phase-valued data.

Tasks. Denoising, Inpainting, labeling (classification), deblurring,...

Manifold-valued Signal & Image Processing

Tasks in image processing are often phrased as an optimisation problem.
Here. The pixel take values on a manifold

- phase-valued data ($\mathbb{S}^1$)
- wind-fields, GPS ($\mathbb{S}^2$)
- DT-MRI ($\mathcal{P}(3)$)
- EBSD, (grain) orientations ($\mathrm{SO}(n)$)



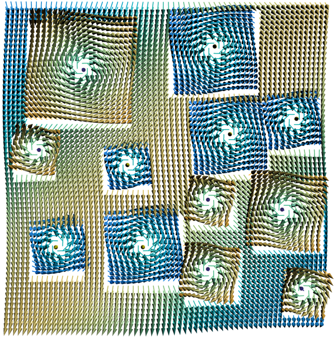
InSAR-Data of Mt. Vesuvius.
[Rocca, Prati, and Guarnieri 1997]

Tasks. Denoising, Inpainting, labeling (classification), deblurring,...

Manifold-valued Signal & Image Processing

Tasks in image processing are often phrased as an optimisation problem.
Here. The pixel take values on a manifold

- phase-valued data ($\mathbb{S}^1$)
- wind-fields, GPS ($\mathbb{S}^2$)
- DT-MRI ($\mathcal{P}(3)$)
- EBSD, (grain) orientations ($\mathrm{SO}(n)$)



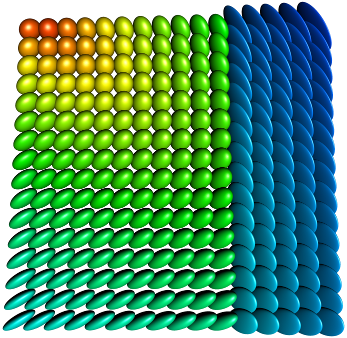
Artificial noisy data on the sphere $\mathbb{S}^2$.

Tasks. Denoising, Inpainting, labeling (classification), deblurring,...

Manifold-valued Signal & Image Processing

Tasks in image processing are often phrased as an optimisation problem.
Here. The pixel take values on a manifold

- phase-valued data ($\mathbb{S}^1$)
- wind-fields, GPS ($\mathbb{S}^2$)
- DT-MRI ($\mathcal{P}(3)$)
- EBSD, (grain) orientations ($\mathrm{SO}(n)$)



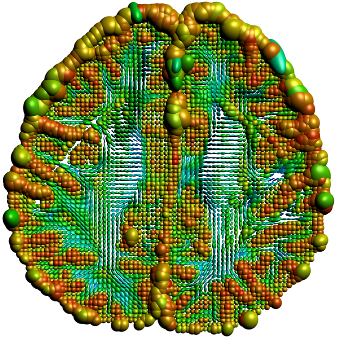
Artificial diffusion data,
each pixel is a symmetric positive matrix.

Tasks. Denoising, Inpainting, labeling (classification), deblurring,...

Manifold-valued Signal & Image Processing

Tasks in image processing are often phrased as an optimisation problem.
Here. The pixel take values on a manifold

- phase-valued data ($\mathbb{S}^1$)
- wind-fields, GPS ($\mathbb{S}^2$)
- DT-MRI ($\mathcal{P}(3)$)
- EBSD, (grain) orientations ($\mathrm{SO}(n)$)



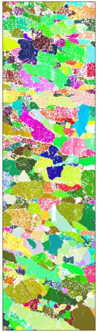
DT-MRI of the human brain.
Camino Profject: cmic.cs.ucl.ac.uk/camino

Tasks. Denoising, Inpainting, labeling (classification), deblurring,...

Manifold-valued Signal & Image Processing

Tasks in image processing are often phrased as an optimisation problem.
Here. The pixel take values on a manifold

- phase-valued data ($\mathbb{S}^1$)
- wind-fields, GPS ($\mathbb{S}^2$)
- DT-MRI ($\mathcal{P}(3)$)
- EBSD, (grain) orientations ($\mathrm{SO}(n)$)



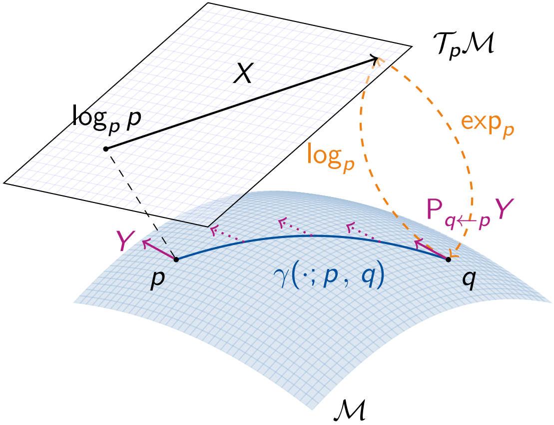
Grain orientations in EBSD data.
MTEX toolbox: mtex-toolbox.github.io

Tasks. Denoising, Inpainting, labeling (classification), deblurring,...

A d-dimensional Riemannian manifold $\mathcal{M}$

Notation.

- Geodesic $\gamma(\cdot; p, q)$
- Tangent space $\mathcal{T}_p\mathcal{M}$
- inner product $(\cdot, \cdot)_p$
- Logarithmic map $\log_p q = \dot\gamma(0; p, q)$
- Exponential map $\exp_p X = \gamma_{p,X}(1)$ where $\gamma_{p,X}(0) = p$ and $\dot\gamma_{p,X}(0) = X$
- Parallel transport $\mathrm{P}_{q\leftarrow p} Y$ "move" tangent vectors from $\mathcal{T}_p\mathcal{M}$ to $\mathcal{T}_q\mathcal{M}$



The Model

We consider a minimization problem

$$\operatorname*{arg\,min}_{p\in\mathcal{C}} F(p) + G(\Lambda(p))$$

- $\mathcal{M}, \mathcal{N}$ are (high-dimensional) Riemannian Manifolds
- $F\colon \mathcal{M} \to \overline{\mathbb{R}}$ nonsmooth, (locally, geodesically) convex
- $G\colon \mathcal{N} \to \overline{\mathbb{R}}$ nonsmooth, (locally) convex
- $\Lambda\colon \mathcal{M} \to \mathcal{N}$ nonlinear
- $\mathcal{C} \subset \mathcal{M}$ strongly geodesically convex.

In image processing.
choose a model, such that finding a minimizer yields the reconstruction

Splitting Methods & Algorithms

On a Riemannian manifold $\mathcal{M}$ we have

- Cyclic Proximal Point Algorithm (CPPA) [Bačák 2014]
- (parallel) Douglas–Rachford Algorithm (PDRA) [RB, Persch, and Steidl 2016]

On $\mathbb{R}^n$ PDRA is known to be equivalent to [Setzer 2011; O'Connor and Vandenberghe 2018]

- Primal-Dual Hybrid Gradient Algorithm (PDHGA) [Esser, Zhang, and Chan 2010]
- Chambolle-Pock Algorithm (CPA) [Chambolle and Pock 2011; Pock, Cremers, Bischof, and Chambolle 2009]

But on a Riemannian manifold $\mathcal{M}$: ⚠ no duality theory!

Goals of this talk.

Formulate Duality on a Manifold

Derive a Riemannian Chambolle–Pock Algorithm (RCPA)

The Euclidean Fenchel Conjugate

Let $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper and convex.
We define the Fenchel conjugate $f^*\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ of f by

$$f^*(\xi) := \sup_{x\in\mathbb{R}^n} \langle \xi, x\rangle - f(x) = \sup_{x\in\mathbb{R}^n} \begin{pmatrix} \xi \\ -1 \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} x \\ f(x) \end{pmatrix}$$

- interpretation: maximize the distance of $\xi^{\mathsf{T}}x$ to f

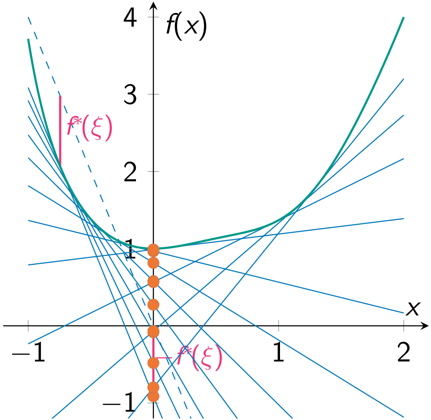
⇒ extremum seeking problem on the epigraph

The Fenchel biconjugate reads

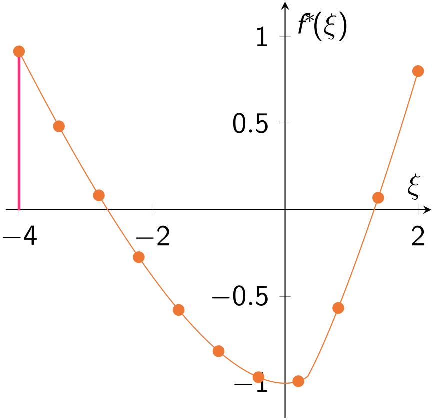
$$f^{**}(x) = (f^*)^*(x) = \sup_{\xi\in\mathbb{R}^n} \langle \xi, x\rangle - f^*(\xi).$$

Illustration of the Fenchel Conjugate

The function f



The Fenchel conjugate f^*



Properties of the Euclidean Fenchel Conjugate

[Rockafellar 1970]

- The Fenchel conjugate f^* is convex (even if f is not)
- f^{**} is the largest convex, lsc function with $f^{**} \leq f$
- If $f(x) \leq g(x)$ holds for all $x \in \mathbb{R}^n$ then $f^*(\xi) \geq g^*(\xi)$ holds for all $\xi \in \mathbb{R}^n$
- Fenchel–Moreau theorem: f convex, proper, lsc $\Rightarrow f^{**} = f$.
- Fenchel–Young inequality:

$$f(x) + f^*(\xi) \geq \xi^\mathsf{T} x \quad \text{for all} \quad x, \xi \in \mathbb{R}^n$$

- For a proper, convex function f

$$\xi \in \partial f(x) \Leftrightarrow f(x) + f^*(\xi) = \xi^\mathsf{T} x$$

- For a proper, convex, lsc function f, then

$$\xi \in \partial f(x) \Leftrightarrow x \in \partial f^*(\xi)$$

The Riemannian $m-$Fenchel Conjugate

[RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]
alternative approaches: [Ahmadi Kakavandi and Amini 2010; Silva Louzeiro, RB, and Herzog 2022]

Idea: Introduce a point on $\mathcal{M}$ to "act as" 0.

Let $m \in \mathcal{C} \subset \mathcal{M}$ be given and $F\colon \mathcal{C} \to \overline{\mathbb{R}}$.
The m-Fenchel conjugate $F_m^*\colon \mathcal{T}_m^*\mathcal{M} \to \overline{\mathbb{R}}$ is defined by

$$F_m^*(\xi_m) := \sup_{X \in \mathcal{L}_{\mathcal{C},m}} \bigl\{ \langle \xi_m , X \rangle - F(\exp_m X) \bigr\},$$

where $\mathcal{L}_{\mathcal{C},m} := \{X \in \mathcal{T}_m\mathcal{M} \mid q = \exp_m X \in \mathcal{C} \text{ and } \|X\|_p = d(q,p)\}$.

Let $m' \in \mathcal{C}$. The mm'-Fenchel-biconjugate $F_{mm'}^{**}\colon \mathcal{C} \to \overline{\mathbb{R}}$ is given by

$$F_{mm'}^{**}(p) = \sup_{\xi_{m'} \in \mathcal{T}_{m'}^*\mathcal{M}} \bigl\{ \langle \xi_{m'} , \log_{m'} p \rangle - F_m^*(\mathrm{P}_{m \leftarrow m'} \xi_{m'}) \bigr\}.$$

usually we only use the case $m = m'$.

Properties of the m-Fenchel Conjugate

[RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]

- F^*_m is convex on $\mathcal{T}^*_m\mathcal{M}$
- $F(p) \leq G(p)$ holds for all $p \in \mathcal{C} \Rightarrow F^*_m(\xi_m) \geq G^*_m(\xi_m)$ holds for all $\xi_m \in \mathcal{T}^*_m\mathcal{M}$
- Fenchel-Moreau theorem: $F \circ \exp_m$ convex (on $\mathcal{T}_m\mathcal{M}$), proper, lsc, then $F^{**}_{mm} = F$ on $\mathcal{C}$.
- Fenchel-Young inequality: For a proper, convex function $F \circ \exp_m$

$$\xi_p \in \partial_{\mathcal{M}} F(p) \Leftrightarrow F(p) + F^*_m(\mathrm{P}_{m\leftarrow p}\xi_p) = \langle \mathrm{P}_{m\leftarrow p}\xi_p \,, \log_m p\rangle.$$

- For a proper, convex, lsc function $F \circ \exp_m$

$$\xi_p \in \partial_{\mathcal{M}} F(p) \Leftrightarrow \log_m p \in \partial F^*_m(\mathrm{P}_{m\leftarrow p}\xi_p).$$

NTNU

Proximal Map

For $F\colon \mathcal{M} \to \overline{\mathbb{R}}$ and $\lambda > 0$ we define the Proximal Map as

[Moreau 1965; Rockafellar 1970; Ferreira and Oliveira 2002]

$$\operatorname{prox}_{\lambda F} p := \operatorname*{arg\,min}_{u \in \mathcal{M}} d(u,p)^2 + \lambda F(u).$$

! For a Minimizer u^* of F we have $\operatorname{prox}_{\lambda F} u^* = u^*$.

- For F proper, convex, lsc:
  - the proximal map is unique.
  - Proximal-Point-Algorithm:
    $x_k = \operatorname{prox}_{\lambda F} x_{k-1}$ converges to $\operatorname{arg\,min} F$
- $q = \operatorname{prox}_{\lambda F} p$ is equivalent to

$$\frac{1}{\lambda}\left(\log_q p\right)^\flat \in \partial_{\mathcal{M}} F(q)$$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

NTNU

[RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]

Input: m, $p^{(0)} \in \mathbb{R}^d$, $n = \Lambda(m)$, $\xi_n^{(0)} \in \mathbb{R}^d$, and parameters σ, τ, $\theta > 0$

$k \leftarrow 0$

$\bar{p}^{(0)} \leftarrow p^{(0)}$

while not converged **do**

$\quad \xi_n^{(k+1)} \leftarrow \operatorname{prox}_{\tau G_n^*}\Bigl(\xi_n^{(k)} + \tau\bigl(\log_n \Lambda(\bar{p}^{(k)})\bigr)^\flat\Bigr)$

$\quad p^{(k+1)} \leftarrow \operatorname{prox}_{\sigma F}\Bigl(p^{(k)} + \mathrm{P}_{p^{(k)} \leftarrow m}\bigl(-\sigma D\Lambda(m)^*[\xi_n^{(k+1)}]\bigr)^\sharp\Bigr)$

$\quad \bar{p}^{(k+1)} \leftarrow p^{(k+1)} + \theta(p^{(k+1)} - p^{(k)})$

$\quad k \leftarrow k + 1$

end while

Output: $p^{(k)}$

Generalizations & Variants of the RCPA

Classically

[Chambolle and Pock 2011]

- change $\sigma = \sigma_k$, $\tau = \tau_k$, $\theta = \theta_k$ during the iterations
- introduce an acceleration γ
- relax dual $\bar{\xi}$ instead of primal $\bar{p}$ (switches lines 4 and 5)

Furthermore we

[RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]

- introduce the IRCPA: linearize Λ, i. e., adopt the Euclidean case from [Valkonen 2014]

$$\log_n \Lambda(\bar{p}^{(k)}) \quad \to \quad \mathrm{P}_{n\leftarrow\Lambda(m)} D\Lambda(m)[\log_m \bar{p}^{(k)}]$$

- choose $n \neq \Lambda(m)$ introduces a parallel transport

$$D\Lambda(m)^*[\xi_n^{(k+1)}] \quad \to \quad D\Lambda(m)^*[\mathrm{P}_{\Lambda(m)\leftarrow n}\xi_n^{(k+1)}]$$

- change $m = m^{(k)}$, $n = n^{(k)}$ during the iterations

Manopt.jl: Optimisation on Manifolds in Julia

[RB 2022]

Goal. Provide optimisation algorithms on Riemannian manifolds, i. e. based on `ManifoldsBase.jl` such that it can be used with all manifolds from `Manifolds.jl`.

Features.

- generic algorithm framework:
  With `Problem P` and `Options O`
  - `initialize_solver!(P,O)`
  - `step_solver!(P, O, i)`: ith step
- run algorithm: call `solve(P,O)`
- generic debug and recording
- step sizes and stopping criteria.

Manopt Family.

- manoptjl.org [RB 2022]
- manopt.org [Boumal, Mishra, Absil, and Sepulchre 2014]
- pymanopt.org [Townsend, Koep, and Weichwald 2016]

Algorithms.

- Gradient Descent
  CG, Stochastic, Momentum, …
- Quasi-Newton
  BFGS, DFP, Broyden, SR1, …
- Nelder-Mead, Particle Swarm
- Subgradient Method
- Trust Regions
- Chambolle-Pock
- Douglas-Rachford
- Cyclic Proximal Point

The ℓ^2-TV Model

[Rudin, Osher, and Fatemi 1992; Lellmann, Strekalovskiy, Koetter, and Cremers 2013; Weinmann, Demaret, and Storath 2014]

For a manifold-valued image $f \in \mathcal{M}$, $\mathcal{M} = \mathcal{N}^{d_1,d_2}$, we compute

$$\operatorname*{arg\,min}_{p\in\mathcal{M}} \frac{1}{\alpha} F(p) + G(\Lambda(p)), \qquad \alpha > 0,$$

with

- data term $F(p) = \frac{1}{2} d^2_{\mathcal{M}}(p, f)$
- "forward differences" $\Lambda\colon \mathcal{M} \to (T\mathcal{M})^{d_1-1,\ d_2-1,\ 2}$,

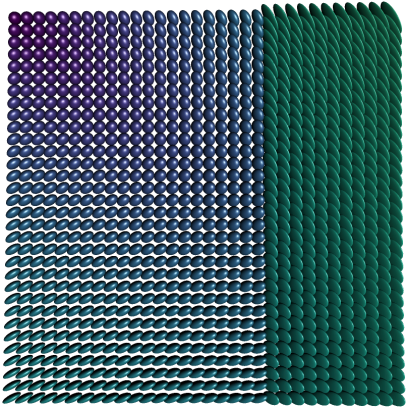
$$p \mapsto \Lambda(p) = \Bigl((\log_{p_i} p_{i+e_1},\ \log_{p_i} p_{i+e_2}) \Bigr)_{i\in\{1,\ldots,d_1-1\}\times\{1,\ldots,d_2-1\}}$$

- prior $G(X) = \|X\|_{g,q,1}$ similar to a collaborative TV [Duran, Moeller, Sbert, and Cremers 2016]

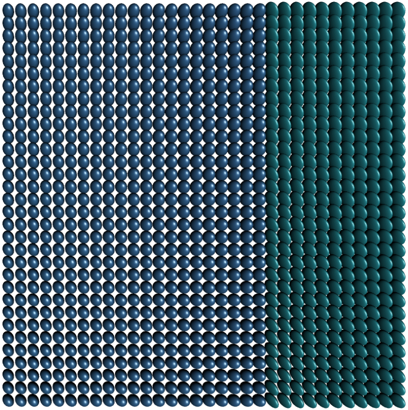
$\Rightarrow$ $\operatorname{prox}_{\lambda G_n^*}$ given in closed form for $q = 1$ (anisotropic TV) and $q = 2$ (isotropic TV).

NTNU

Numerical Example for a $\mathcal{P}(3)$-valued Image



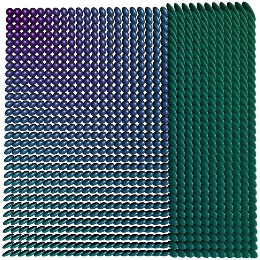
$\mathcal{P}(3)$-valued data.



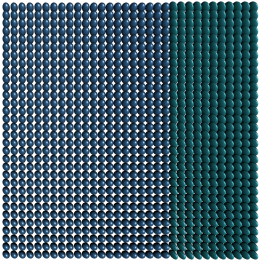
anisotropic TV, $\alpha = 6$.

- in each pixel we have a symmetric positive definite matrix
- Applications: denoising/inpainting e.g. of DT-MRI data

Numerical Example for a $\mathcal{P}(3)$-valued Image



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

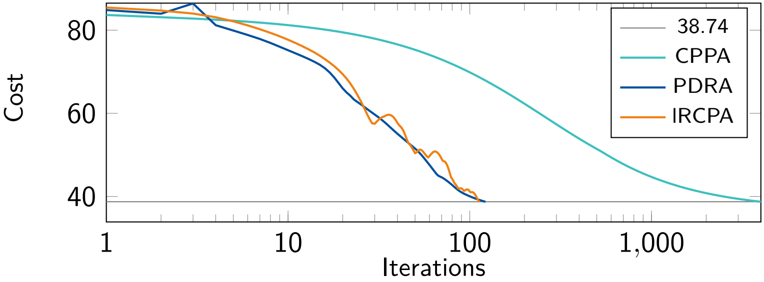


anisotropic TV, $\alpha = 6$.

Approach. CPPA as benchmark [Bačák 2014; RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]

	CPPA	PDRA	IRCPA
parameters	$\lambda_k = \frac{4}{k}$	$\lambda = 0.58$ $\beta = 0.93$	$\sigma = \tau = 0.4$ $\gamma = 0.2,\ m = I$
iterations	4000	122	**113**
runtime	1235 s.	380 s.	**96.1 s.**

Numerical Example for a $\mathcal{P}(3)$-valued Image



Approach. CPPA as benchmark [Bačák 2014; RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]

	CPPA	PDRA	IRCPA
parameters	$\lambda_k = \frac{4}{k}$	$\lambda = 0.58$ $\beta = 0.93$	$\sigma = \tau = 0.4$ $\gamma = 0.2,\ m = I$
iterations	4000	122	**113**
runtime	1235 s.	380 s.	**96.1 s.**

Summary

Summary.

- We introduced a duality framework on manifolds
- we introduced a Riemannian Chambolle–Pock algorithm
- We saw a Software framework for Optimisation algorithms on manifolds
- Numerical examples illustrates its performance

Further.

- A Fenchel Conjugate can be defined on $T\mathcal{M}$
  $\Rightarrow$ (hyperplane) separation Theorem [RB, Herzog, and Silva Louzeiro 2021]
- Extended to semi-smooth Newrton [Diepeveen and Lellmann 2021]

Outlook.

- Strategies for choosing base points, investigate $C(k)$
- Investigate constraint optimisation on Manifolds
- look into further applications

Selected References

- Axen, S. D., M. Baran, RB, and K. Rzecki (2021). *Manifolds.jl: An Extensible Julia Framework for Data Analysis on Manifolds*. arXiv: 2106.08777.
- RB (2022). "Manopt.jl: Optimization on Manifolds in Julia". In: *Journal of Open Source Software* 7.70, p. 3866. DOI: 10.21105/joss.03866.
- RB, R. Herzog, and M. Silva Louzeiro (2021). *Fenchel duality and a separation theorem on Hadamard manifolds*. arXiv: 2102.11155.
- RB, R. Herzog, M. Silva Louzeiro, D. Tenbrinck, and J. Vidal-Núñez (Jan. 2021). "Fenchel duality theory and a primal-dual algorithm on Riemannian manifolds". In: *Foundations of Computational Mathematics*. DOI: 10.1007/s10208-020-09486-5. arXiv: 1908.02022.
- RB, J. Persch, and G. Steidl (2016). "A parallel Douglas Rachford algorithm for minimizing ROF-like functionals on images with values in symmetric Hadamard manifolds". In: *SIAM Journal on Imaging Sciences* 9.4, pp. 901–937. DOI: 10.1137/15M1052858.
- Chambolle, A. and T. Pock (2011). "A first-order primal-dual algorithm for convex problems with applications to imaging". In: *Journal of Mathematical Imaging and Vision* 40.1, pp. 120–145. DOI: 10.1007/s10851-010-0251-1.
- Diepeveen, W. and J. Lellmann (2021). "An Inexact Semismooth Newton Method on Riemannian Manifolds with Application to Duality-Based Total Variation Denoising". In: *SIAM Journal on Imaging Sciences* 14.4, pp. 1565–1600. DOI: 10.1137/21m1398513.

ronnybergmann.net/talks/2022-GAMM-Chambolle-Pock.pdf