

# The Riemannian Chambolle-Pock algorithm

Ronny Bergmann

joint work with

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

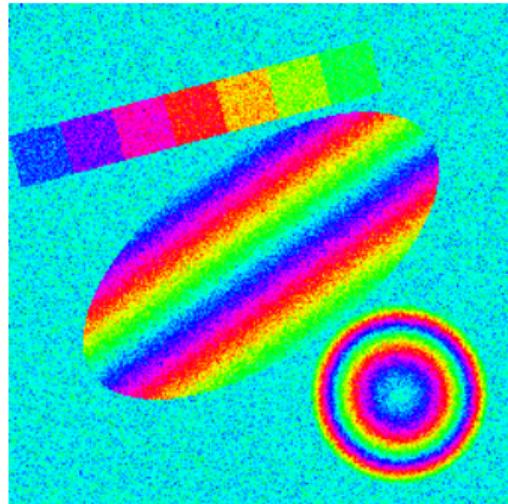
Jahrestagung der Gesellschaft für angewandte Mathematik und Mechanik,  
Session 21. Mathematical signal and image processing

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 ( $\text{SO}(n)$ )



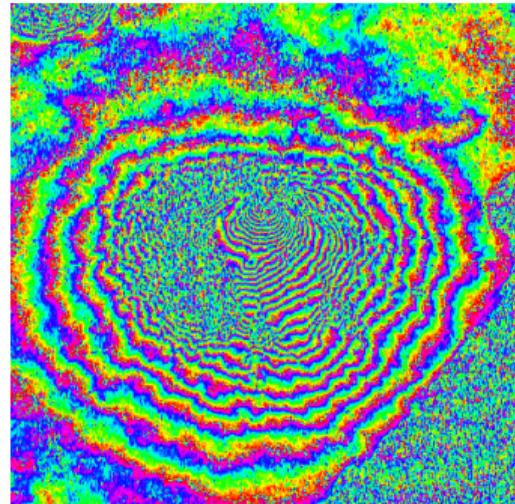
Artificial noisy phase-valued data.

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

# Manifold-valued Signal & Image Processing

Tasks in [image processing](#) are often phrased as an optimisation problem.  
**Here.** The pixels 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 ( $\text{SO}(n)$ )



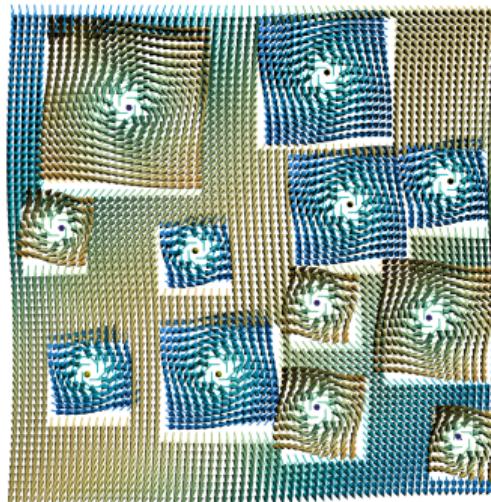
InSAR-Data of Mt. Vesuvius.  
[Rocca, Prati, and Guarnieri 1997]

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

# Manifold-valued Signal & Image Processing

Tasks in [image processing](#) are often phrased as an optimisation problem.  
**Here.** The pixels 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 ( $\text{SO}(n)$ )



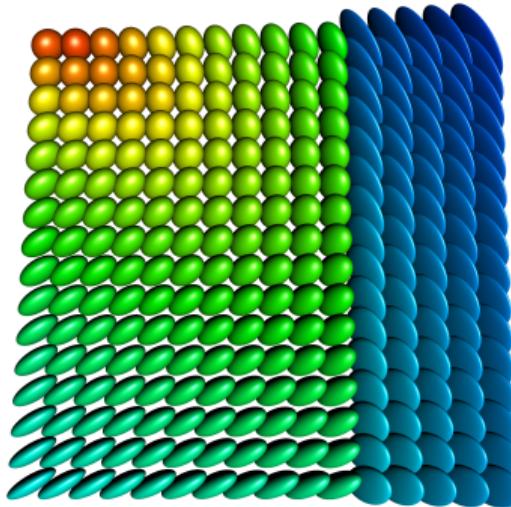
Artificial noisy data on the sphere  $\mathbb{S}^2$ .

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

# Manifold-valued Signal & Image Processing

Tasks in [image processing](#) are often phrased as an optimisation problem.  
**Here.** The pixels 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 ( $\text{SO}(n)$ )



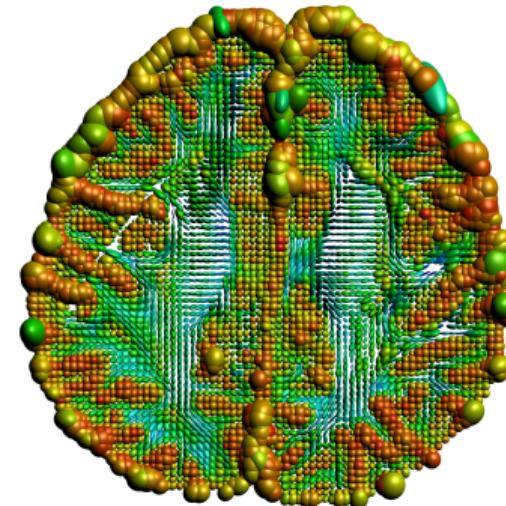
Artificial diffusion data,  
each pixel is a symmetric positive matrix.

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

# Manifold-valued Signal & Image Processing

Tasks in [image processing](#) are often phrased as an optimisation problem.  
**Here.** The pixels 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 ( $\text{SO}(n)$ )



DT-MRI of the human brain.

Camino Project: [cmic.cs.ucl.ac.uk/camino](http://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 ( $\text{SO}(n)$ )



Grain orientations in EBSD data.

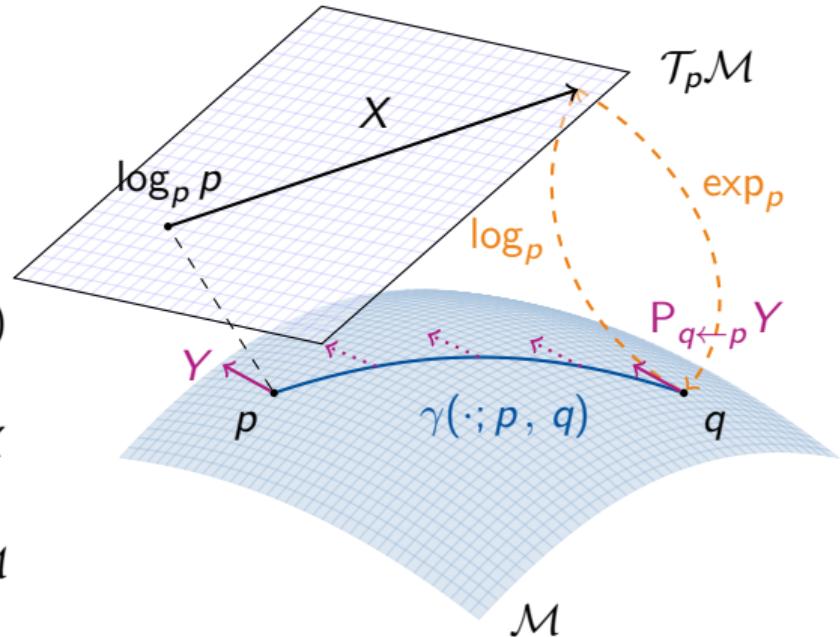
MTEX toolbox: [mTEX toolbox](https://mtex-toolbox.github.io).

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

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

## Notation.

- ▶ Geodesic  $\gamma(\cdot; p, q)$
- ▶ Tangent space  $T_p\mathcal{M}$
- ▶ 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  $P_{q \leftarrow p} Y$  “move”  
tangent vectors from  $T_p\mathcal{M}$  to  $T_q\mathcal{M}$





# The Model

NTNU

We consider a minimization problem

$$\arg \min_{p \in \mathcal{C}} F(p) + G(\Lambda(p))$$

- ▶  $\mathcal{M}, \mathcal{N}$  are (high-dimensional) Riemannian Manifolds
- ▶  $F: \mathcal{M} \rightarrow \overline{\mathbb{R}}$  nonsmooth, (locally, geodesically) convex
- ▶  $G: \mathcal{N} \rightarrow \overline{\mathbb{R}}$  nonsmooth, (locally) convex
- ▶  $\Lambda: \mathcal{M} \rightarrow \mathcal{N}$  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)
- ▶ (parallel) Douglas–Rachford Algorithm (PDRA)

[Bačák 2014]

[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)
- ▶ Chambolle-Pock Algorithm (CPA)

[Esser, Zhang, and Chan 2010]

[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

NTNU

Let  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be proper and convex.

We define the **Fenchel conjugate**  $f^*: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  of  $f$  by

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

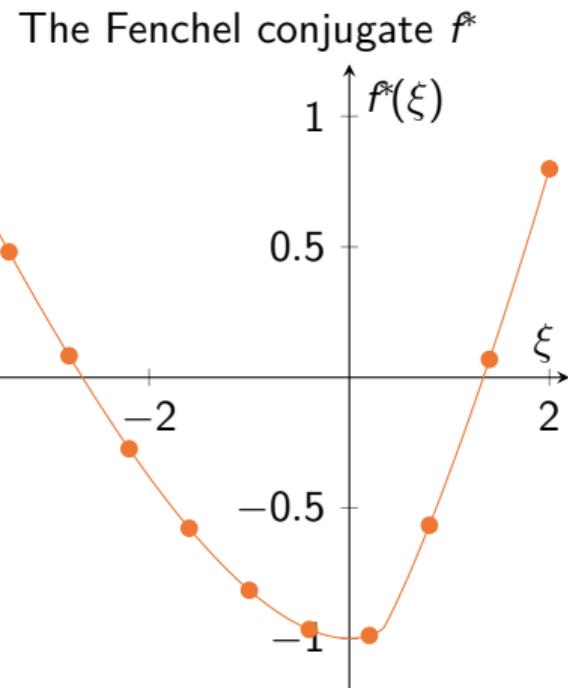
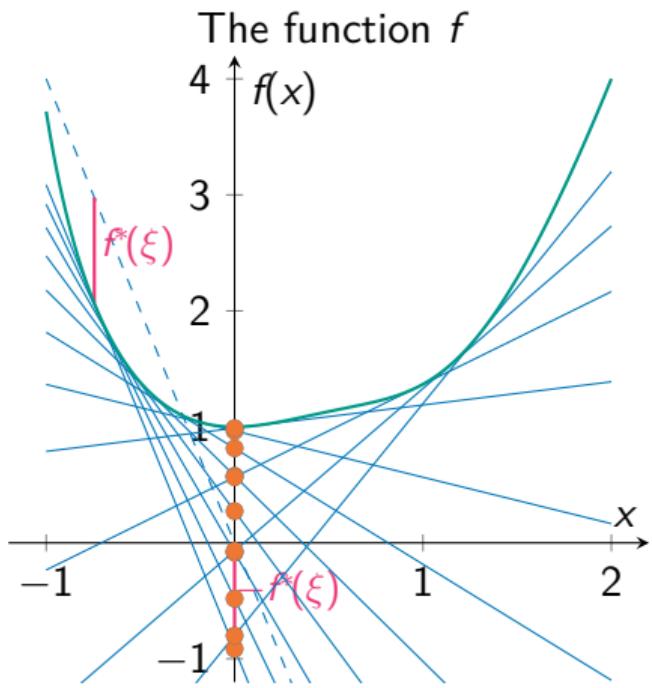
► interpretation: maximize the distance of  $\xi^\top x$  to  $f$

⇒ 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 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  $F: \mathcal{C} \rightarrow \overline{\mathbb{R}}$ .

The  $m$ -Fenchel conjugate  $F_m^*: \mathcal{T}_m^*\mathcal{M} \rightarrow \overline{\mathbb{R}}$  is defined by

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

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  $F_{mm'}^{**}: \mathcal{C} \rightarrow \overline{\mathbb{R}}$  is given by

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

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  $F \circ \exp_m$

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

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

$$\xi_p \in \partial_{\mathcal{M}} F(p) \Leftrightarrow \log_m p \in \partial F_m^*(P_{m \leftarrow p} \xi_p).$$



# Proximal Map

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

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

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

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

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

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



# 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  $\sigma, \tau, \theta > 0$

1:  $k \leftarrow 0$

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

3: **while** not converged **do**

4:      $\xi_n^{(k+1)} \leftarrow \text{prox}_{\tau G_n^*}(\xi_n^{(k)} + \tau (\log_n \Lambda(\bar{p}^{(k)}))^\flat)$

5:      $p^{(k+1)} \leftarrow \text{prox}_{\sigma F}\left(p^{(k)} + P_{p^{(k)} \leftarrow m}(-\sigma D\Lambda(m)^*[\xi_n^{(k+1)}])^\sharp\right)$

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

7:      $k \leftarrow k + 1$

8: **end while**

**Output:**  $p^{(k)}$

# Generalizations & Variants of the RCPA



NTNU

Classically

[Chambolle and Pock 2011]

- ▶ change  $\sigma = \sigma_k$ ,  $\tau = \tau_k$ ,  $\theta = \theta_k$  during the iterations
- ▶ introduce an acceleration  $\gamma$
- ▶ 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  $\Lambda$ , i. e., adopt the Euclidean case from  
[Valkonen 2014]

$$\log_n \Lambda(\bar{p}^{(k)}) \rightarrow P_{n \leftarrow \Lambda(m)} D\Lambda(m) [\log_m \bar{p}^{(k)}]$$

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

$$D\Lambda(m)^* [\xi_n^{(k+1)}] \rightarrow D\Lambda(m)^* [P_{\Lambda(m) \leftarrow n} \xi_n^{(k+1)}]$$

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



[RB 2022]

# Manopt.jl: Optimisation on Manifolds in Julia



NTNU

**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)`: *i*th 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]

## Algoirthms.

- ▶ 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 $\ell^2$ -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

$$\arg \min_{p \in \mathcal{M}} \frac{1}{\alpha} F(p) + G(\Lambda(p)), \quad \alpha > 0,$$

with

- ▶ data term  $F(p) = \frac{1}{2} d_{\mathcal{M}}^2(p, f)$
- ▶ “forward differences”  $\Lambda: \mathcal{M} \rightarrow (T\mathcal{M})^{d_1-1, d_2-1, 2}$ ,

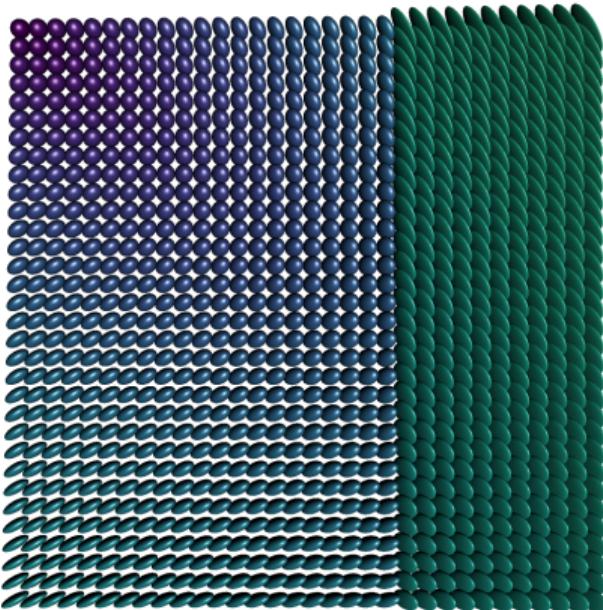
$$p \mapsto \Lambda(p) = \left( (\log_{p_i} p_{i+e_1}, \log_{p_i} p_{i+e_2}) \right)_{i \in \{1, \dots, d_1-1\} \times \{1, \dots, d_2-1\}}$$

- ▶ prior  $G(X) = \|X\|_{g,q,1}$  similar to a collaborative TV [Duran, Moeller, Sbert, and Cremers 2016]
- ⇒ prox $_{\lambda G_n^*}$  given in closed form for  $q = 1$  (anisotropic TV) and  $q = 2$  (isotropic TV).

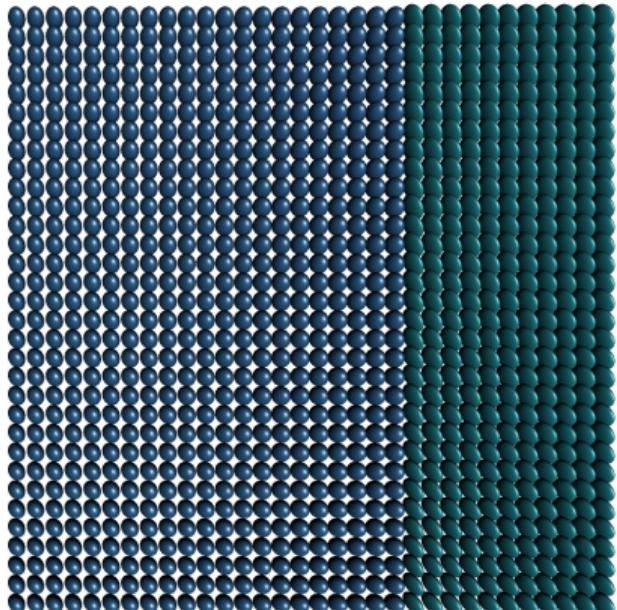
# Numerical Example for a $\mathcal{P}(3)$ -valued Image



NTNU



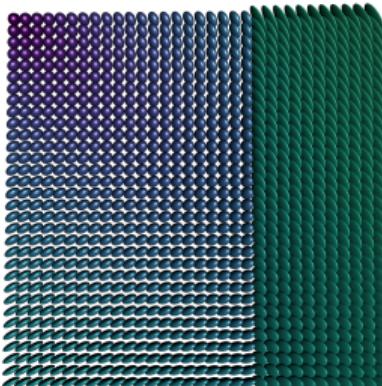
$\mathcal{P}(3)$ -valued data.



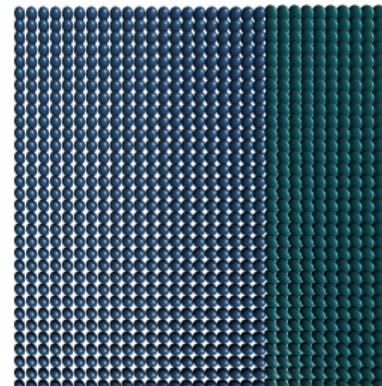
anisotropic TV,  $\alpha = 6$ .

- ▶ 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.



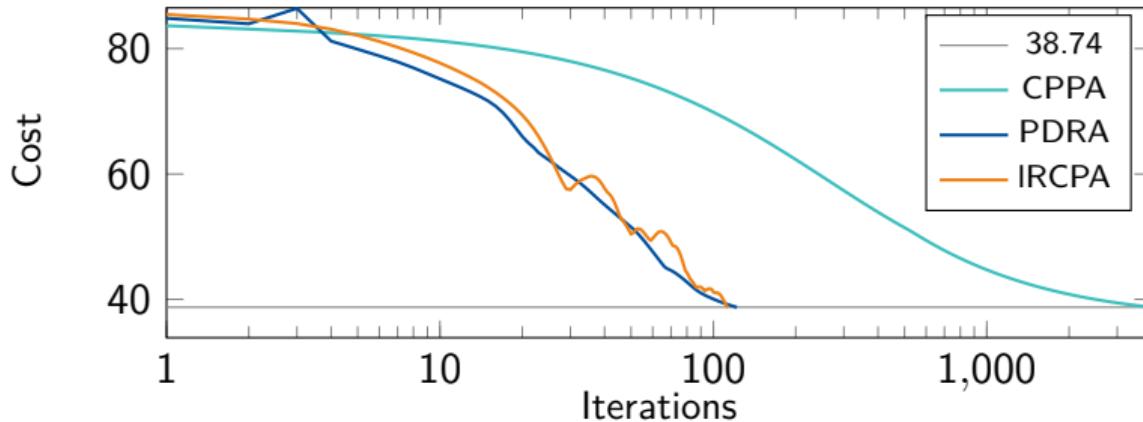
anisotropic TV,  $\alpha = 6$ .

**Approach.** CPPA as benchmark

[Bačák 2014; RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]

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

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



**Approach.** CPPA as benchmark [Bacák 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	<b>113</b>
runtime	1235 s.	380 s.	<b>96.1 s.</b>



# Summary

NTNU

## 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

## 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.
-  Bačák, M. (2014). “Computing medians and means in Hadamard spaces”. In: *SIAM Journal on Optimization* 24.3, pp. 1542–1566. DOI: 10.1137/140953393.
-  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.