The Riemannian Chambolle–Pock Algorithm

Ronny Bergmann

joint work with
Roland Herzog, Maurício Silva Louzeiro, Daniel Tenbrinck, José Vidal-Núñez.

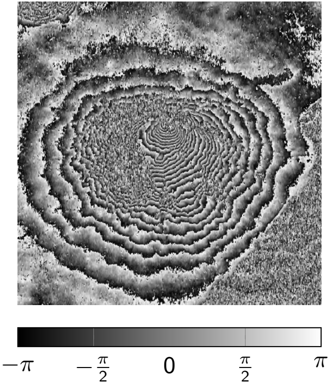
Manifolds and Geometric Integration Colloquia, Ilsetra, March 3, 2022

Norwegian University of Science and Technology

Manifold-Valued Signals and Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals,
  GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI),
  covariance matrices
- Electron backscattered diffraction (EBSD)



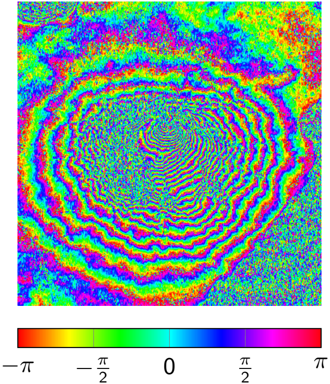
InSAR-Data of Mt. Vesuvius
[Rocca, Prati, and Guarnieri 1997]

phase-valued data, $\mathcal{M} = \mathbb{S}^1$

Manifold-Valued Signals and Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals,
  GPS data, wind, flow,…
- Diffusion tensors in magnetic resonance imaging (DT-MRI),
  covariance matrices
- Electron backscattered diffraction (EBSD)



InSAR-Data of Mt. Vesuvius

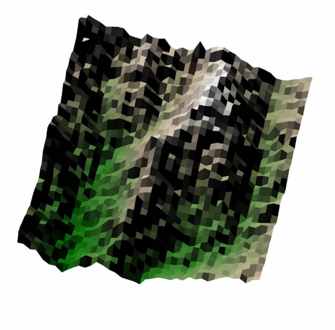
[Rocca, Prati, and Guarnieri 1997]

phase-valued data, $\mathcal{M} = \mathbb{S}^1$

Manifold-Valued Signals and Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)

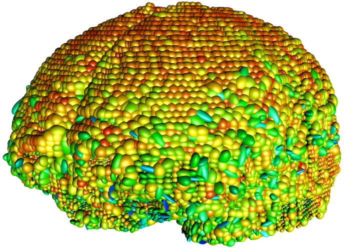


National elevation dataset
[Gesch, Evans, Mauck, Hutchinson, and Carswell Jr 2009]
directional data, $\mathcal{M} = \mathbb{S}^2$

Manifold-Valued Signals and Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



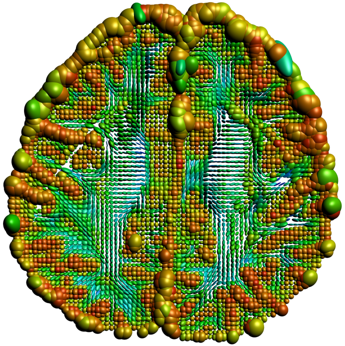
diffusion tensors in human brain from the Camino dataset
http://cmic.cs.ucl.ac.uk/camino

sym. pos. def. matrices, $\mathcal{M} = \mathrm{SPD}(3)$

Manifold-Valued Signals and Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,…
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



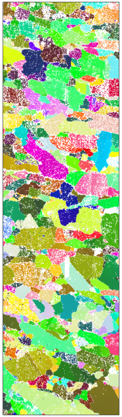
horizontal slice # 28
from the Camino dataset
http://cmic.cs.ucl.ac.uk/camino

sym. pos. def. matrices, $\mathcal{M} = \mathrm{SPD}(3)$

Manifold-Valued Signals and Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



EBSD example from the MTEX toolbox
Bachmann and Hielscher, since 2007

Rotations (mod. symmetry), $\mathcal{M} = \mathrm{SO}(3)(/\mathcal{S})$.

NTNU

Manifold-Valued Signals and Images

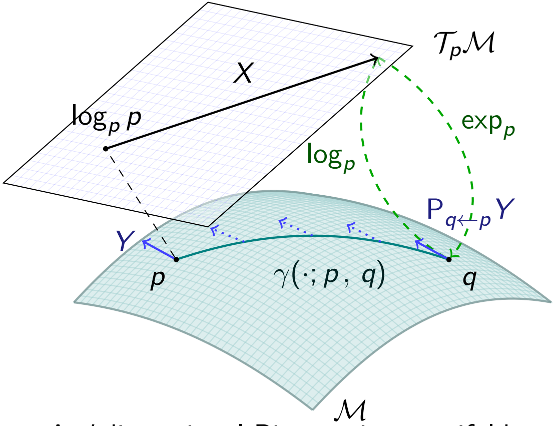
New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals,
  GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI),
  covariance matrices
- Electron backscattered diffraction (EBSD)

Common properties

- Range of values is a Riemannian manifold
- Tasks from "classical" image processing, e.g.
  - denoising
  - inpainting
  - interpolation
  - labeling
  - deblurring

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

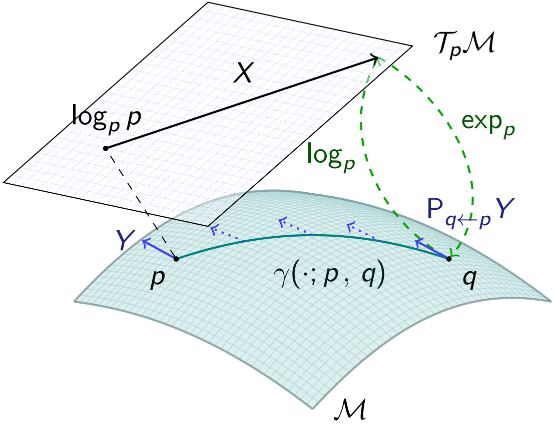


A d-dimensional Riemannian manifold can be informally defined as a set $\mathcal{M}$ covered with a 'suitable' collection of charts, that identify subsets of $\mathcal{M}$ with open subsets of $\mathbb{R}^d$ and a continuously varying inner product on the tangent spaces.

[Absil, Mahony, and Sepulchre 2008]

NTNU

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



Geodesic $\gamma(\cdot; p, q)$

a shortest path between $p, q \in \mathcal{M}$

Tangent space $\mathcal{T}_p\mathcal{M}$ at p

with inner product $(\cdot, \cdot)_p$

Logarithmic map $\log_p q = \dot\gamma(0; p, q)$

"speed towards 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 $\mathrm{P}_{q\leftarrow p}Y$

from $\mathcal{T}_p\mathcal{M}$ along $\gamma(\cdot; p, q)$ 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 [O'Connor and Vandenberghe 2018; Setzer 2011]

- 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

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

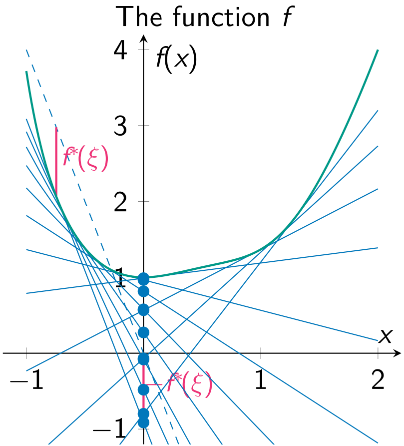
- interpretation: maximize the distance of $\xi^{\mathrm{T}} x$ to f

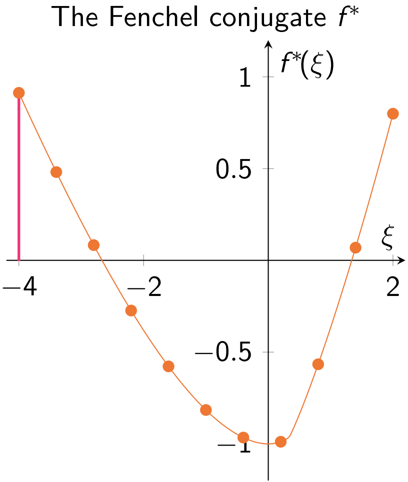
$\Rightarrow$ extremum seeking problem on the epigraph

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





NTNU

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

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 $m = m'$.

The Euclidean Chambolle–Pock Algorithm

[Chambolle and Pock 2011]

From the pair of primal-dual problems

$$\min_{x\in\mathbb{R}^n} f(x) + g(Kx), \quad K \text{ linear},$$
$$\max_{\xi\in\mathbb{R}^m} -f^*(-K^*\xi) - g^*(\xi)$$

we obtain for f, g proper convex, lsc the
optimality conditions (OC) for a solution $(\hat{x}, \hat{\xi})$ as

$$\begin{aligned} \partial f &\ni -K^*\hat{\xi} \\ \partial g^*(\hat{\xi}) &\ni K\hat{x} \end{aligned}$$

The Euclidean Chambolle–Pock Algorithm

[Chambolle and Pock 2011]

From the pair of primal-dual problems

$$\min_{x\in\mathbb{R}^n} f(x) + g(Kx), \quad K \text{ linear},$$
$$\max_{\xi\in\mathbb{R}^m} - f^*(-K^*\xi) - g^*(\xi)$$

we obtain for f, g proper convex, lsc the

,

Chambolle–Pock Algorithm. with $\sigma > 0$, $\tau > 0$, $\theta \in \mathbb{R}$ reads

$$\begin{aligned}
x^{(k+1)} &= \operatorname{prox}_{\sigma f}\bigl(x^{(k)} - \sigma K^*\bar\xi^{(k)}\bigr)\\
\xi^{(k+1)} &= \operatorname{prox}_{\tau g^*}\bigl(\xi^{(k)} + \tau K x^{(k+1)}\bigr)\\
\bar\xi^{(k+1)} &= \xi^{(k+1)} + \theta\bigl(\xi^{(k+1)} - \xi^{(k)}\bigr)
\end{aligned}$$

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.
  - PPA $x_k = \operatorname{prox}_{\lambda F} x_{k-1}$ converges to $\operatorname{arg\,min} F$
- $q = \operatorname{prox}_{\lambda F} p$ is equivalent to

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

Saddle Point Formulation

Let F be geodesically convex, $G \circ \exp_n$ be convex (on $\mathcal{T}_n\mathcal{N}$).

From

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

we derive the saddle point formulation for the n-Fenchel conjugate of G as

$$\min_{p \in \mathcal{C}} \max_{\xi_n \in \mathcal{T}_n^*\mathcal{N}} \langle \xi_n \,, \log_n \Lambda(p) \rangle + F(p) - G_n^*(\xi_n).$$

But $\Lambda\colon \mathcal{M} \to \mathcal{N}$ is a non-linear operator!

For Optimality Conditions and the Dual Prolem: What's Λ^*?

[Valkonen 2014]

Approach. Linearization:

$$\Lambda(p) \approx \exp_{\Lambda(m)} D\Lambda(m)[\log_m p]$$

The exact Riemannian Chambolle–Pock Algorithm (eRCPA)

Input: $m, p^{(0)} \in \mathcal{C} \subset \mathcal{M}$, $n = \Lambda(m)$, $\xi_n^{(0)} \in \mathcal{T}_n^*\mathcal{N}$,
and parameters σ, τ, $\theta > 0$

1: $k \leftarrow 0$
2: $\bar{p}^{(0)} \leftarrow p^{(0)}$
3: **while** not converged **do**
4: $\xi_n^{(k+1)} \leftarrow \operatorname{prox}_{\tau G_n^*} \xi_n^{(k)} + \tau \big(\log_n \Lambda(\bar{p}^{(k)}) \big)^\flat$
5: $p^{(k+1)} \leftarrow \operatorname{prox}_{\sigma F} \exp_{p^{(k)}} \Big(\mathrm{P}_{m \leftarrow p^{(k)}} \big(-\sigma D\Lambda(m)^*[\xi_n^{(k+1)}] \big)^\sharp \Big)$
6: $\bar{p}^{(k+1)} \leftarrow \exp_{p^{(k+1)}} \big(-\theta \log_{p^{(k+1)}} p^{(k)} \big)$
7: $k \leftarrow k + 1$
8: **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

Manifolds.jl: A Library of Manifolds in Julia

[Axen, Baran, RB, and Rzecki 2021]

`ManifoldsBase.jl` provides a unified interface to implement & use manifolds
also provides e.g. `ValidationManifold` (for debugging) and an `EmbeddedManifold`.

`Manifolds.jl` uses this interface to provide

Features.

- different metrics
- Lie groups
- Build manifolds using
  - Product manifold $\mathcal{M}_1 \times \mathcal{M}_2$
  - Power manifold $\mathcal{M}^{n \times m}$
  - Tangent bundle
- perform statistics

Manifolds. For example

- (unit) Sphere
- Circle & Torus
- Fixed Rank Matrices
- Stiefel & Grassmann
- Hyperbolic space
- Rotations
- Symmetric positive definite matrices
- Symplectic & Symplectic Stiefel
- ...

see `https://juliamanifolds.github.io/Manifolds.jl/`

Manopt.jl: Optimization on Manifolds in Julia

Build upon `ManifoldsBase.jl` to solve [RB 2022]

$$\operatorname*{arg\,min}_{q\in\mathcal{M}} f(q)$$

using

- a `Problem p` describing function, gradient, Hessian,...
- `Options o` specifying a solver settings and state
- call `solve(p, o)`, which includes `StoppingCriterion` calls

implement your own solver within the solver framework

- `initialize_solver!(p, o)` to set up the solver
- `step_solver!(p, o, i)` to perform the ith step

The Manopt family: manoptjl.org

`Manopt` in Matlab [N. Boumal] manopt.org

`pymanopt` in Python [J. Townsend, N. Koep, S. Weichwald] pymanopt.org

and similar: GeomStats (Python), ROPTLIB (C++)

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

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

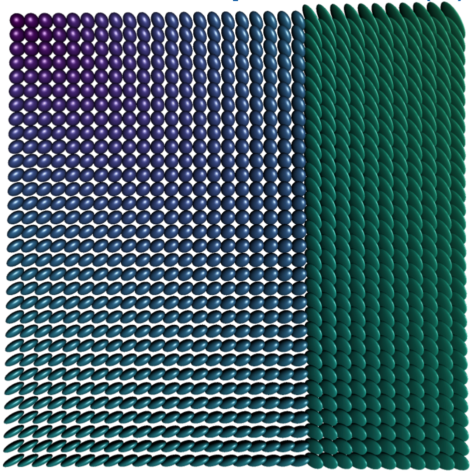
with

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

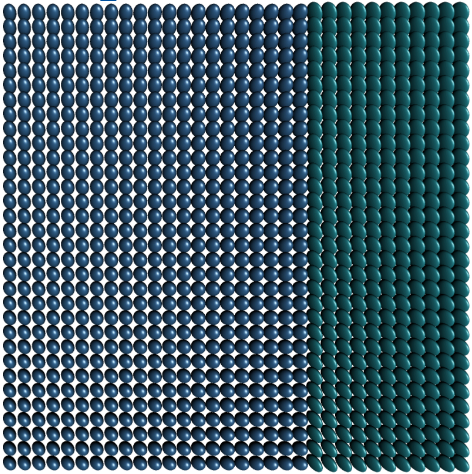
$$p \mapsto \Lambda(p) = \Big(\big(\log_{p_i} p_{i+e_1},\ \log_{p_i} p_{i+e_2}\big) \Big)_{i\in\{1,\ldots,d_1-1\}\times\{1,\ldots,d_2-1\}}$$

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

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



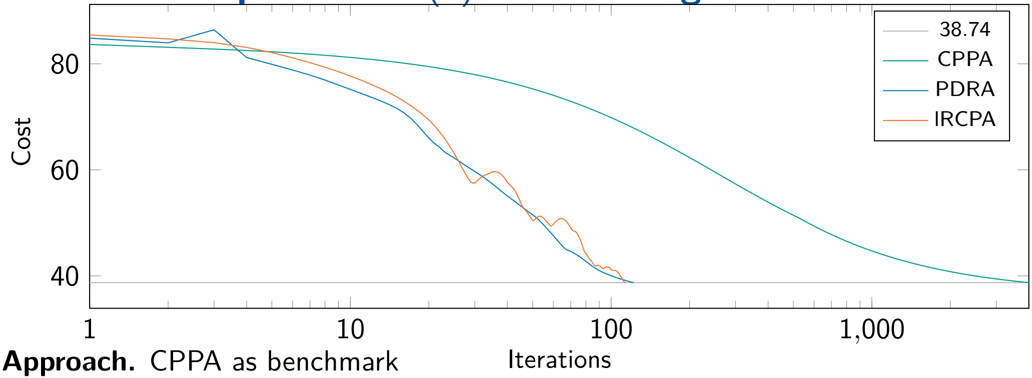
$\mathcal{P}(3)$-valued data.



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



Approach. CPPA as benchmark

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

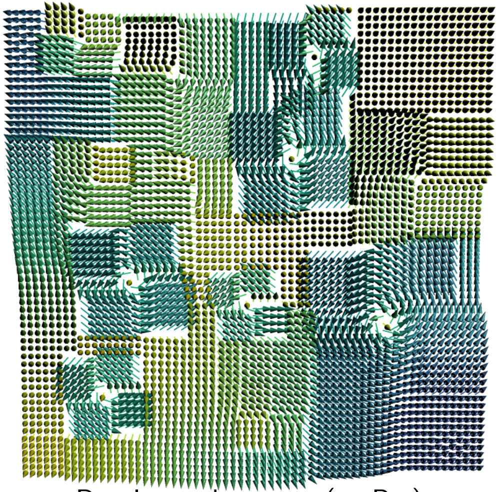
NTNU

Base point Effect on $\mathbb{S}^2$-valued data

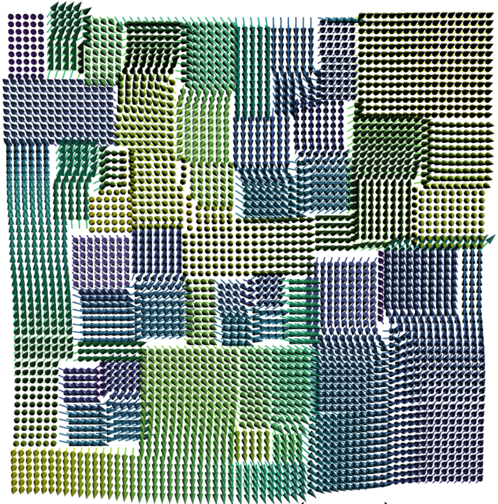
Original data

Original data

Base point Effect on $\mathbb{S}^2$-valued data

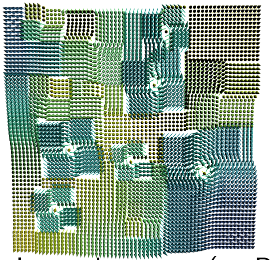


Result, m the mean (p. Px.)

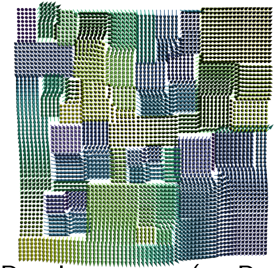


Result, m west (p. Px.)

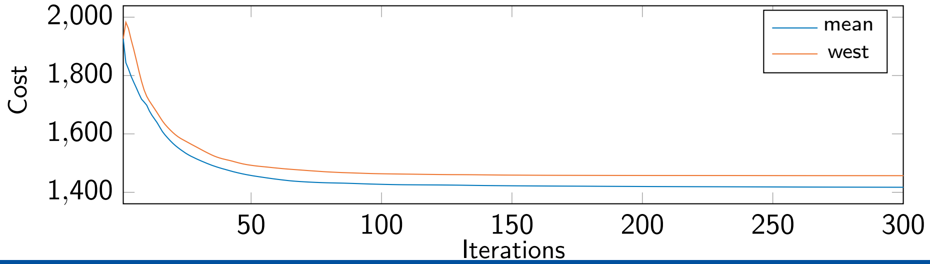
Base point Effect on $\mathbb{S}^2$-valued data



Result, m the mean (p. Px.)



Result, m west (p. Px.)



Summary & Outlook

Summary.

- We introduced a duality framework
  on Riemannian manifolds
- We derived a Riemannian Chambolle Pock Algorithm
- Numerical example illustrates performance

Outlook.

- strategies for choosing m, n (adaptively)
- investigate linearization error
- We started a package `ManifoldsDiffEq.jl`
  https://github.com/JuliaManifolds/ManifoldDiffEq.jl
  to combine `OrdinaryDiffEq.jl` and `ManifoldsBase.jl`

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, 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.
- Silva Louzeiro, M., RB, and R. Herzog (2022). *Fenchel Duality and a Separation Theorem on Hadamard Manifolds*. accepted for publication. arXiv: 2102.11155.

ronnybergmann.net/talks/2022-MaGIC-RiemannianChambollePock.pdf