A Primal-Dual Algorithm for Convex Nonsmooth Optimization on Riemannian Manifolds

Ronny Bergmann

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

INFORMS Annual Meeting 2021, Session “Optimization on Manifolds”
Annaheim, CA, USA & virtual, October 24–27, 2021

Norwegian University of Science and Technology

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!

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]

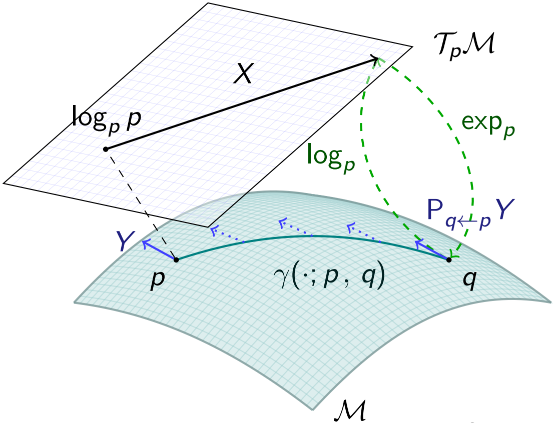
But on a Riemannian manifold $\mathcal{M}$: ⚠ no duality theory!

Goals of this talk.

Formulate Duality on a Manifold

Derive a Riemannian Chambolle–Pock Algorithm (RCPA)

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

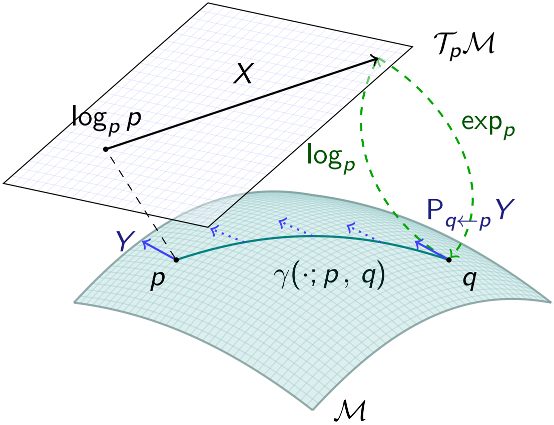


A d-dimensional Riemannian manifold can be informally defined as a set $\mathcal{M}$ covered with a 'suitable' collection of charts, that identify subsets of $\mathcal{M}$ with open subsets of $\mathbb{R}^d$ and a continuously varying inner product on the tangent spaces.

[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 Euclidean Fenchel Conjugate

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

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

The Euclidean Fenchel Conjugate

We define the Fenchel conjugate $f^*\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ of $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ 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

$\Rightarrow$ extremum seeking problem on the epigraph

The Euclidean Fenchel Conjugate

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

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

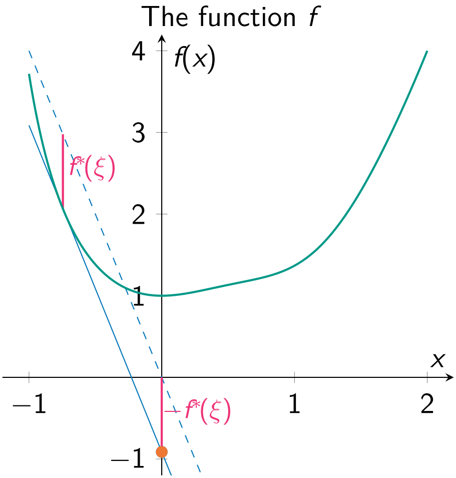
- interpretation: maximize the distance of $\xi^{\mathrm{T}} x$ to f

$\Rightarrow$ 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



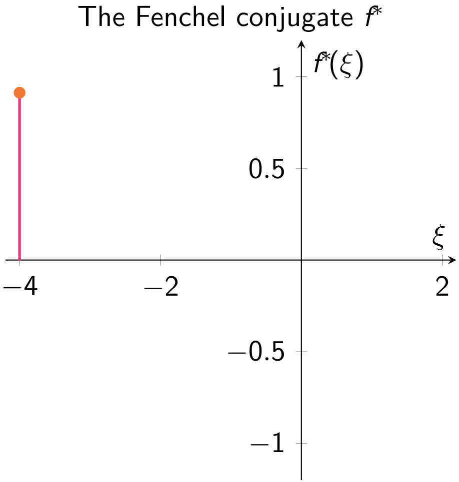


Illustration of the Fenchel Conjugate

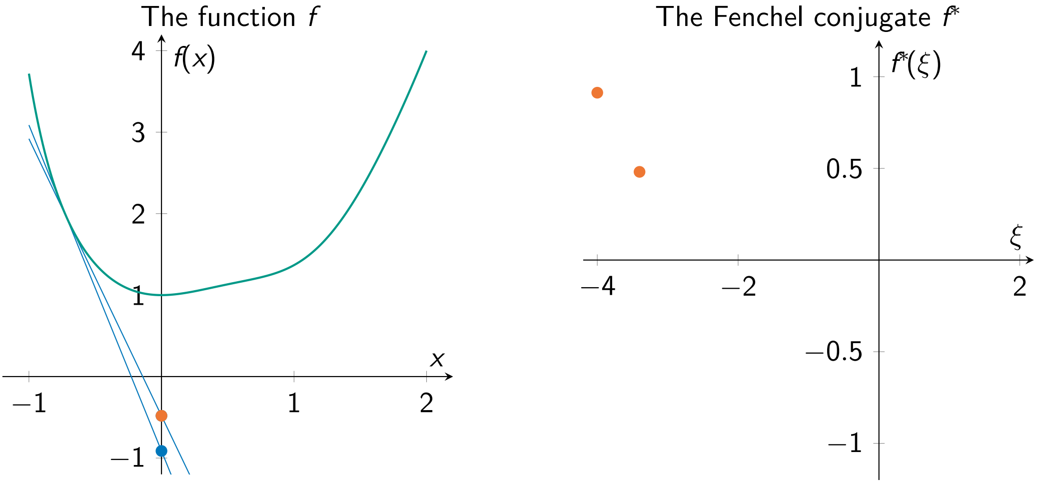
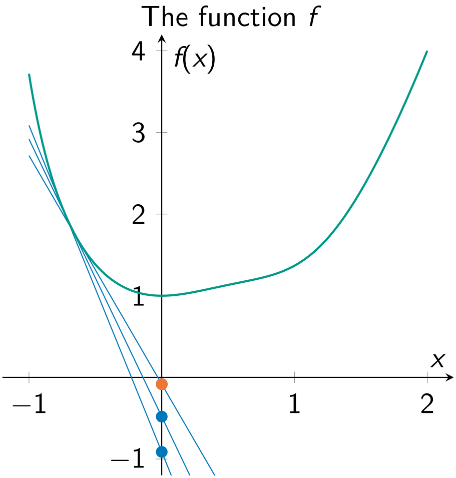


Illustration of the Fenchel Conjugate



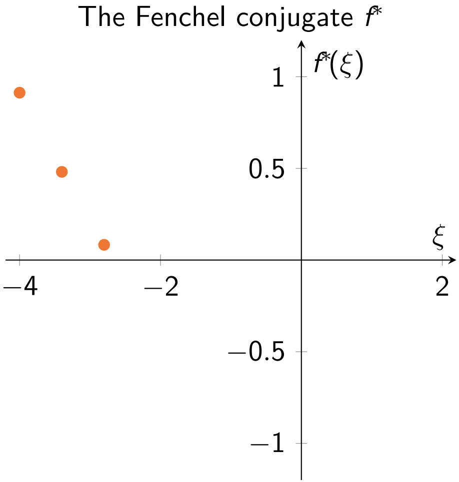
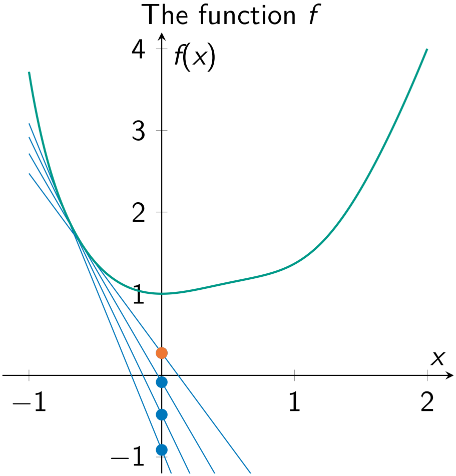
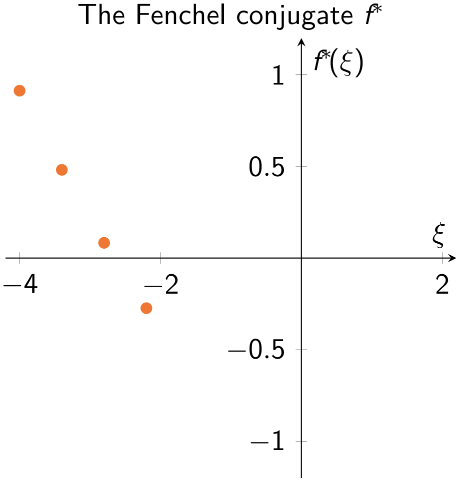


Illustration of the Fenchel Conjugate





NTNU

Illustration of the Fenchel Conjugate

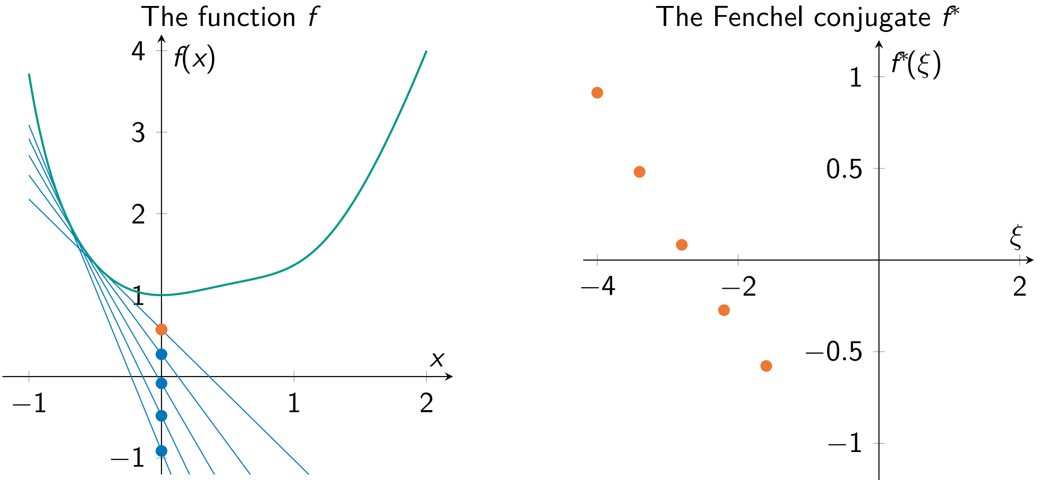
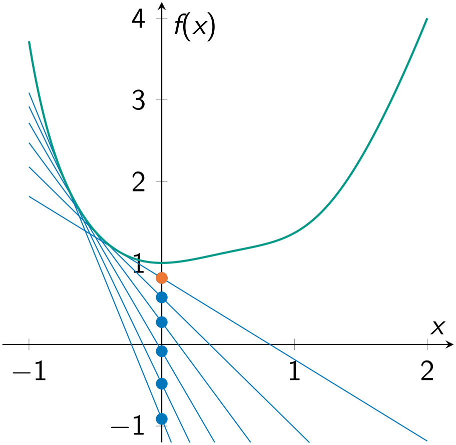
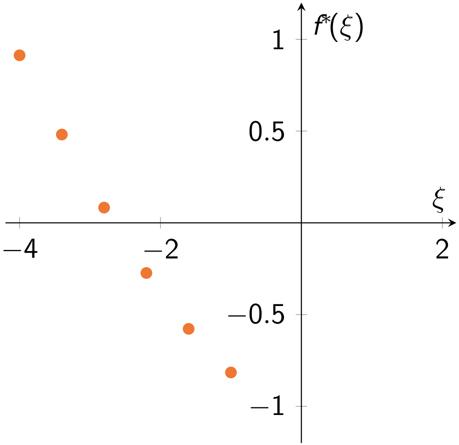


Illustration of the Fenchel Conjugate

The function f

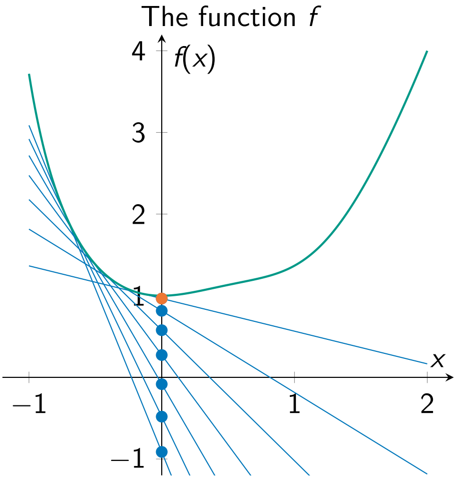


The Fenchel conjugate f^*



NTNU

Illustration of the Fenchel Conjugate



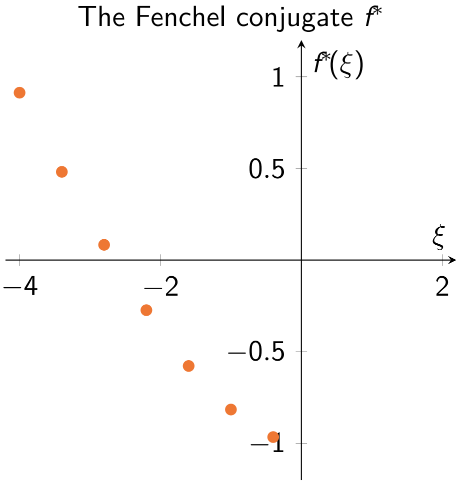
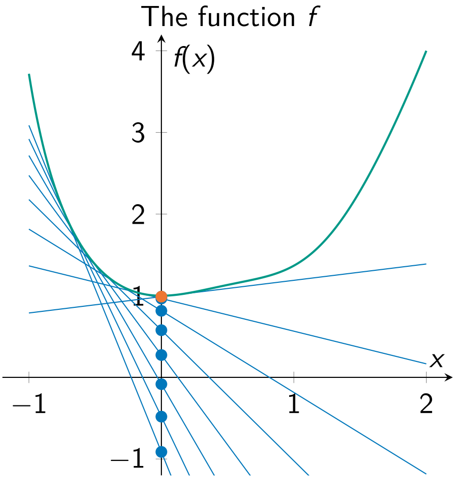


Illustration of the Fenchel Conjugate



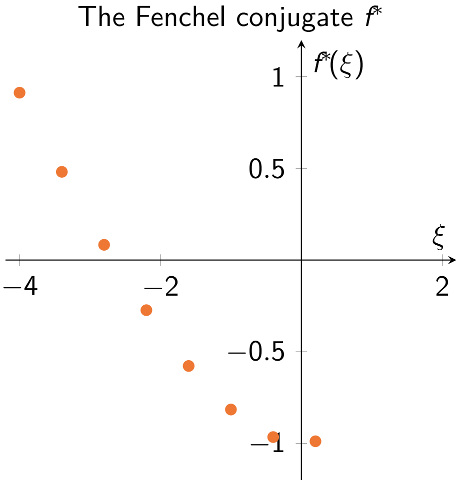
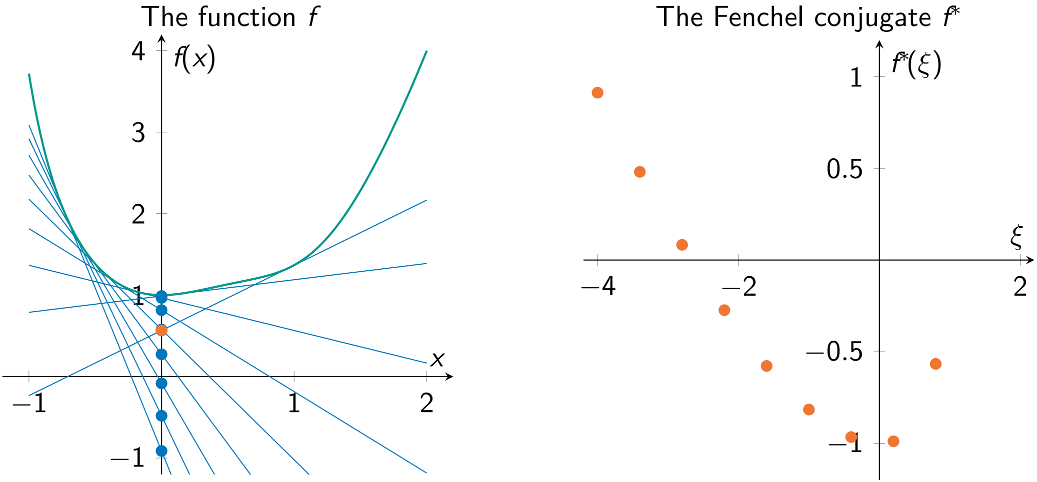
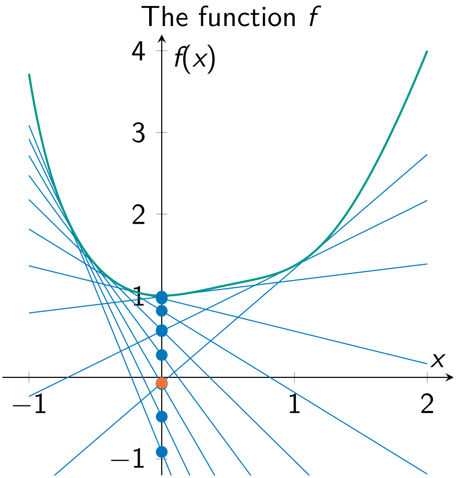


Illustration of the Fenchel Conjugate



NTNU

Illustration of the Fenchel Conjugate



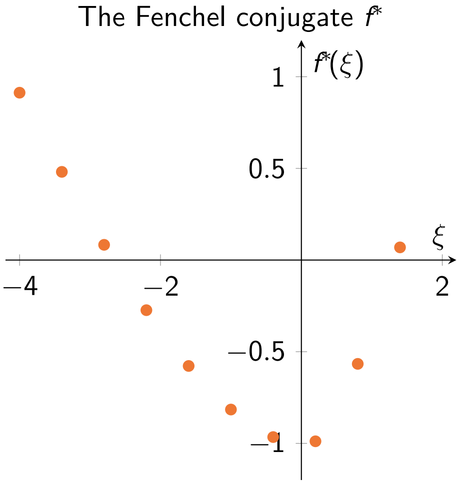
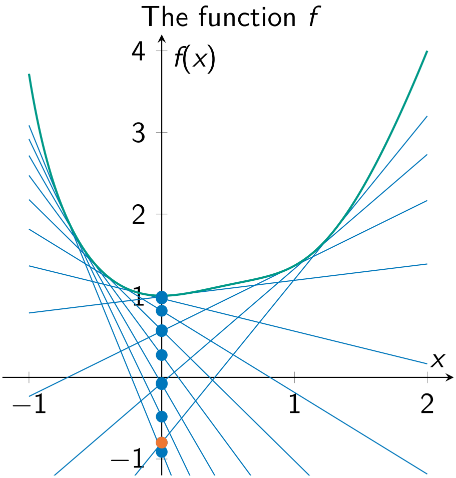


Illustration of the Fenchel Conjugate



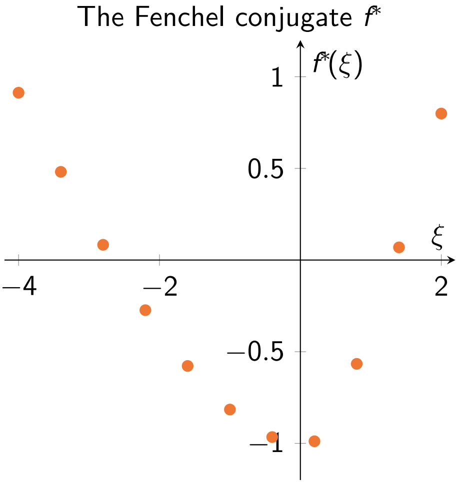
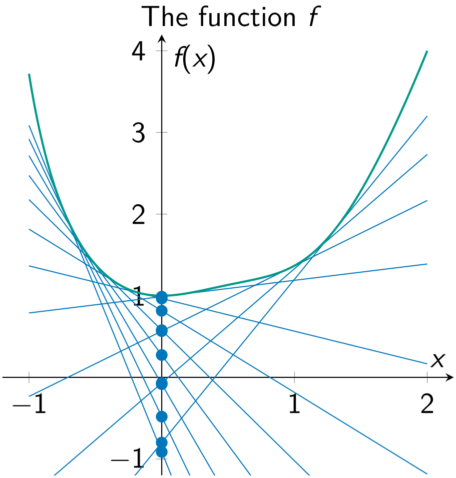
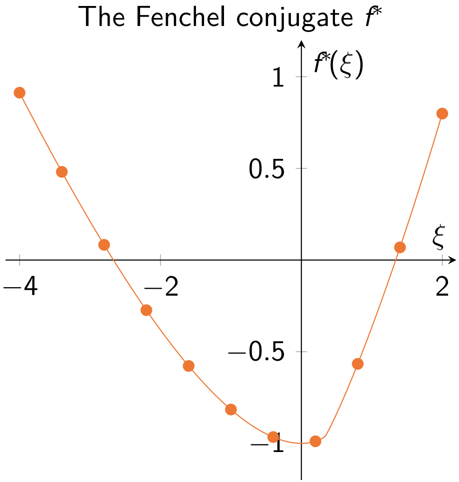


Illustration of the Fenchel Conjugate





The Riemannian $m-$Fenchel Conjugate

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

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

The Riemannian $m-$Fenchel Conjugate

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

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

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

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

The Riemannian $m-$Fenchel Conjugate

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

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

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

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

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 $m = m'$.

NTNU

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^*?

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

NTNU

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

Input: $p^{(0)} \in \mathbb{R}^d$, $\xi^{(0)} \in \mathbb{R}^d$, and parameters σ, τ, $\theta > 0$

1: $k \leftarrow 0$
2: $\bar{p}^{(0)} \leftarrow p^{(0)}$
3: **while** not converged **do**
4: $\quad \xi^{(k+1)} \leftarrow \operatorname{prox}_{\tau G^*}\big(\xi^{(k)} + \tau(\quad \Lambda(\bar{p}^{(k)})) \ \big)$
5: $\quad p^{(k+1)} \leftarrow \operatorname{prox}_{\sigma F}\Big(p^{(k)} \qquad \big(-\sigma \ \Lambda \quad {}^* \ \xi^{(k+1)} \ \big)^\sharp \qquad \Big)$
6: $\quad \bar{p}^{(k+1)} \leftarrow p^{(k+1)} + \theta(p^{(k+1)} - p^{(k)})$
7: $\quad k \leftarrow k+1$
8: **end while**

Output: $p^{(k)}$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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

1: $k \leftarrow 0$

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

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

4: $\xi^{(k+1)} \leftarrow \operatorname{prox}_{\tau G^*}\bigl(\xi^{(k)} + \tau(\quad \Lambda(\bar{p}^{(k)})) \ \bigr)$

5: $p^{(k+1)} \leftarrow \operatorname{prox}_{\sigma F}\Bigl(p^{(k)} \qquad (-\sigma \ \ \Lambda \quad {}^* \xi^{(k+1)})^\sharp \quad \Bigr)$

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

7: $k \leftarrow k + 1$

8: **end while**

Output: $p^{(k)}$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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, $\theta > 0$

$k \leftarrow 0$

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

while not converged **do**

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

$\quad p^{(k+1)} \leftarrow \operatorname{prox}_{\sigma F}\Bigl(p^{(k)} \qquad \bigl(-\sigma \ \Lambda \quad {}^{*}\ \xi_n^{(k+1)}\ \bigr)^{\sharp} \qquad \Bigr)$

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

$\quad k \leftarrow k + 1$

end while

Output: $p^{(k)}$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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

5: $\quad p^{(k+1)} \leftarrow \operatorname{prox}_{\sigma F}\Bigl(p^{(k)} \qquad \bigl(-\sigma \quad \Lambda \qquad {}^* \; \xi_n^{(k+1)}\bigr)^\sharp \qquad \Bigr)$

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

7: $\quad k \leftarrow k + 1$

8: **end while**

Output: $p^{(k)}$

NTNU

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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

1: $k \leftarrow 0$

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

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

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

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

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

7: $\quad k \leftarrow k + 1$

8: **end while**

Output: $p^{(k)}$

NTNU

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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

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

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

7: $\quad k \leftarrow k + 1$

8: **end while**

Output: $p^{(k)}$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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

1: $k \leftarrow 0$

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

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

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

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

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

7: $\qquad k \leftarrow k+1$

8: **end while**

Output: $p^{(k)}$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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, $\theta > 0$

$k \leftarrow 0$

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

while not converged **do**

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

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

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

$\quad k \leftarrow k + 1$

end while

Output: $p^{(k)}$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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

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

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

7: $\quad k \leftarrow k + 1$

8: **end while**

Output: $p^{(k)}$

The Exact Riemannian Chambolle–Pock Algorithm (eRCPA)

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

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: $\quad \xi_n^{(k+1)} \leftarrow \operatorname{prox}_{\tau G_n^*}\Big(\xi_n^{(k)} + \tau\big(\log_n \Lambda(\bar{p}^{(k)})\big)^\flat\Big)$

5: $\quad p^{(k+1)} \leftarrow \operatorname{prox}_{\sigma F}\Bigg(\exp_{p^{(k)}}\Big(\mathrm{P}_{p^{(k)} \leftarrow m}\big(-\sigma D\Lambda(m)^*[\xi_n^{(k+1)}]\big)^\sharp\Big)\Bigg)$

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

7: $\quad k \leftarrow k + 1$

8: **end while**

Output: $p^{(k)}$

NTNU

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

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

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

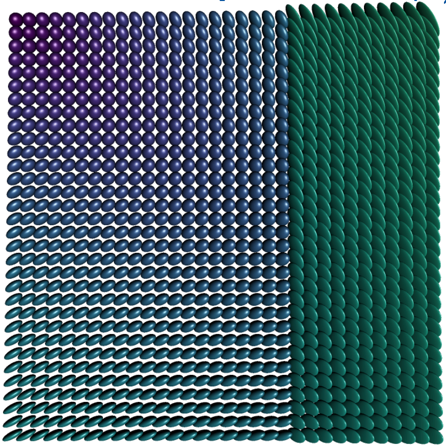
with

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

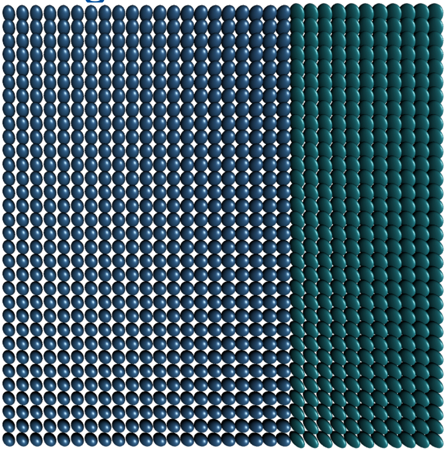
$$p \mapsto \Lambda(p) = \Big((\log_{p_i} p_{i+e_1},\ \log_{p_i} p_{i+e_2}) \Big)_{i \in \{1,\ldots,d_1-1\} \times \{1,\ldots,d_2-1\}}$$

- 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



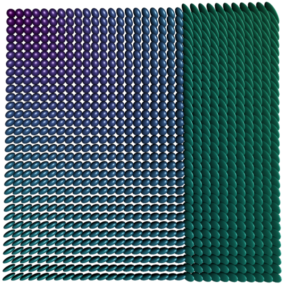
$\mathcal{P}(3)$-valued data.



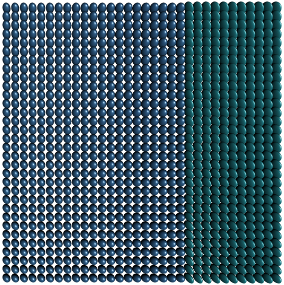
anisotropic TV, $\alpha = 6$.

- 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



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

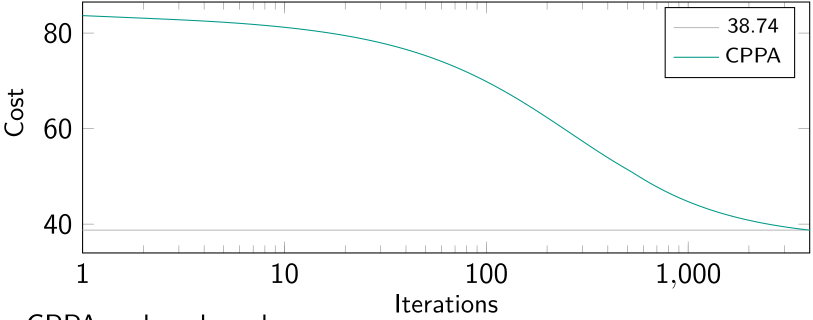


anisotropic TV, $\alpha = 6$.

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		
runtime	1235 s.		

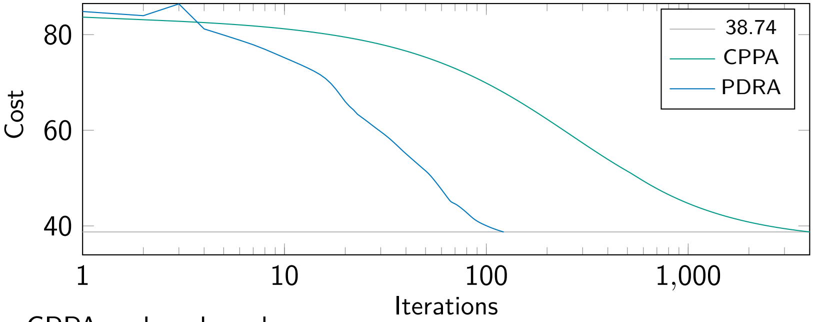
Numerical Example for a $\mathcal{P}(3)$-valued Image



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		
runtime	1235 s.		

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

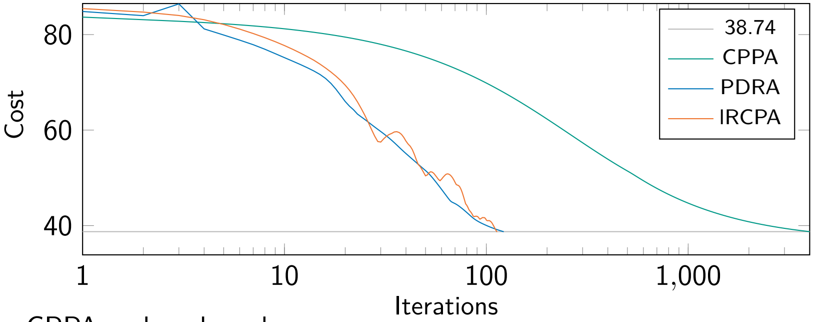


[Bačák 2014; RB, Herzog, Silva Louzeiro, Tenbrinck, and Vidal-Núñez 2021]

Approach. CPPA as benchmark

	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	
runtime	1235 s.	380 s.	

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



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

Approach. CPPA as benchmark

	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 & Outlook

Summary.

- We introduced a duality framework on Riemannian manifolds
- We derived a Riemannian Chambolle Pock Algorithm
- Numerical examples illustrate performance

Outlook.

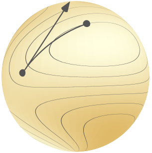
- investigate $C(k)$ and the error of linearization
- strategies for choosing m, n (adaptively)
- alternative models of Fenchel duality (e. g. without m) [RB, Herzog, and Silva Louzeiro 2021]
- higher order methods non-smooth methods [Diepeveen and Lellmann 2021]

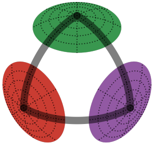
NTNU

Reproducible Research

The algorithm is published in `Manopt.jl`, a Julia Package available at http://manoptjl.org.

It uses the interface from `ManifoldsBase.jl` and hence any manifold from `Manifolds.jl` can be used in the algorithms.
https://juliamanifolds.github.io/Manifolds.jl/
[Axen, Baran, RB, and Rzecki 2021]





Goal.

Being able to use an(y) algorithm for a(ny) model directly on a(ny) manifold easily and efficiently.

Alternatives.

- Manopt, manopt.org (Matlab, by N. Boumal)
- pymanopt, pymanopt.github.io (Python, by S. Weichwald, J. Townsend, N. Koep)

Selected References

- Ahmadi Kakavandi, B. and M. Amini (Nov. 2010). "Duality and subdifferential for convex functions on complete metric spaces". In: *Nonlinear Analysis: Theory, Methods & Applications* 73.10, pp. 3450–3455. DOI: 10.1016/j.na.2010.07.033.
- 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, 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.
- 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.
- Valkonen, T. (2014). "A primal–dual hybrid gradient method for nonlinear operators with applications to MRI". In: *Inverse Problems* 30.5, p. 055012. DOI: 10.1088/0266-5611/30/5/055012.

ronnybergmann.net/talks/2021-INFORMS-RCPA.pdf