

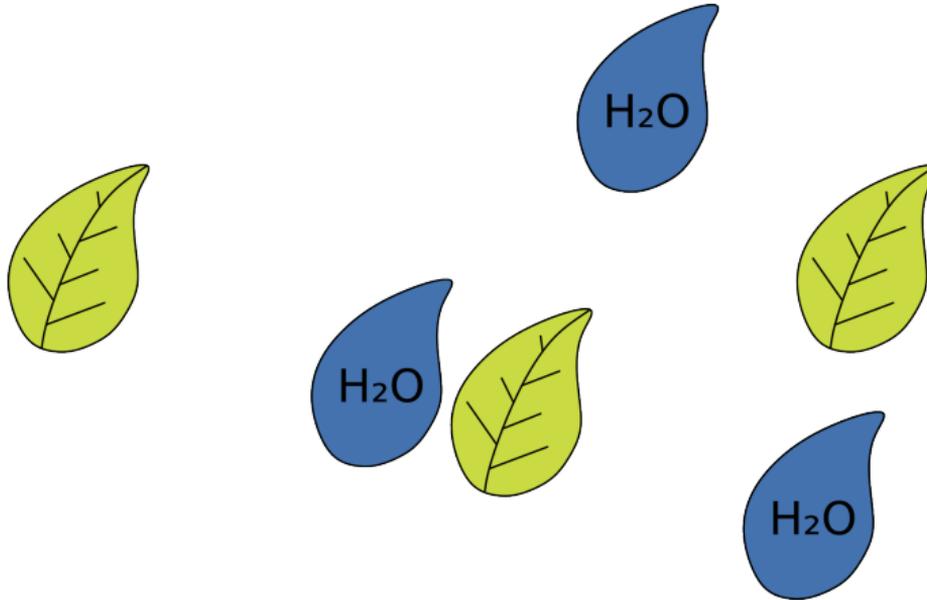
Sliced Optimal Transport on the Sphere

Michael Quellmalz | TU Berlin | CSIP 2023
Chemnitz Symposium on Inverse Problems, Würzburg, 8 November 2023
Joint work with Robert Beinert and Gabriele Steidl

Content

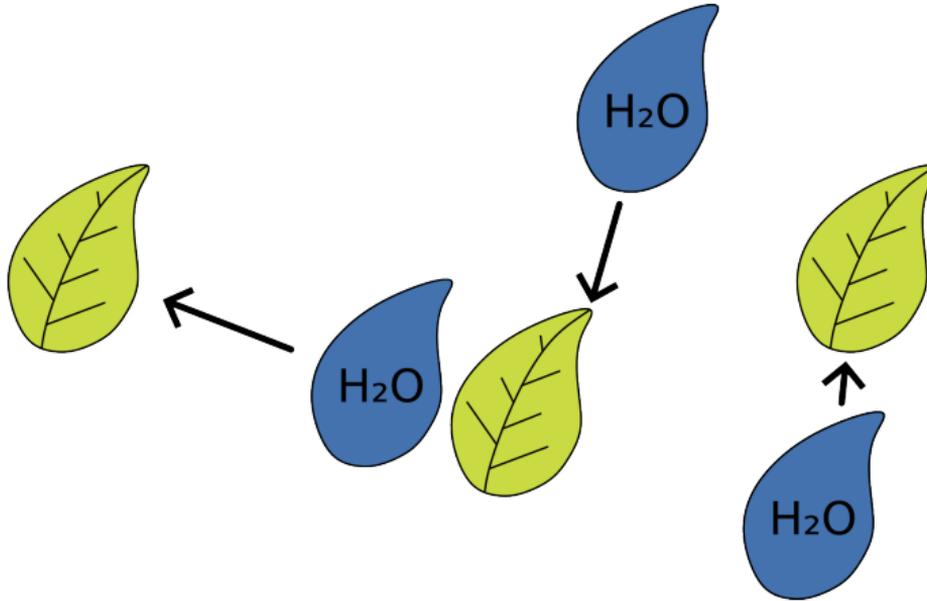
- 1 Introduction to Optimal Transport**
- 2 Sliced and 1-dimensional Optimal Transport
- 3 Sliced Optimal Transport on the Sphere
- 4 Numerics

Introductory Example: Water Your Plants Optimal Transport Style

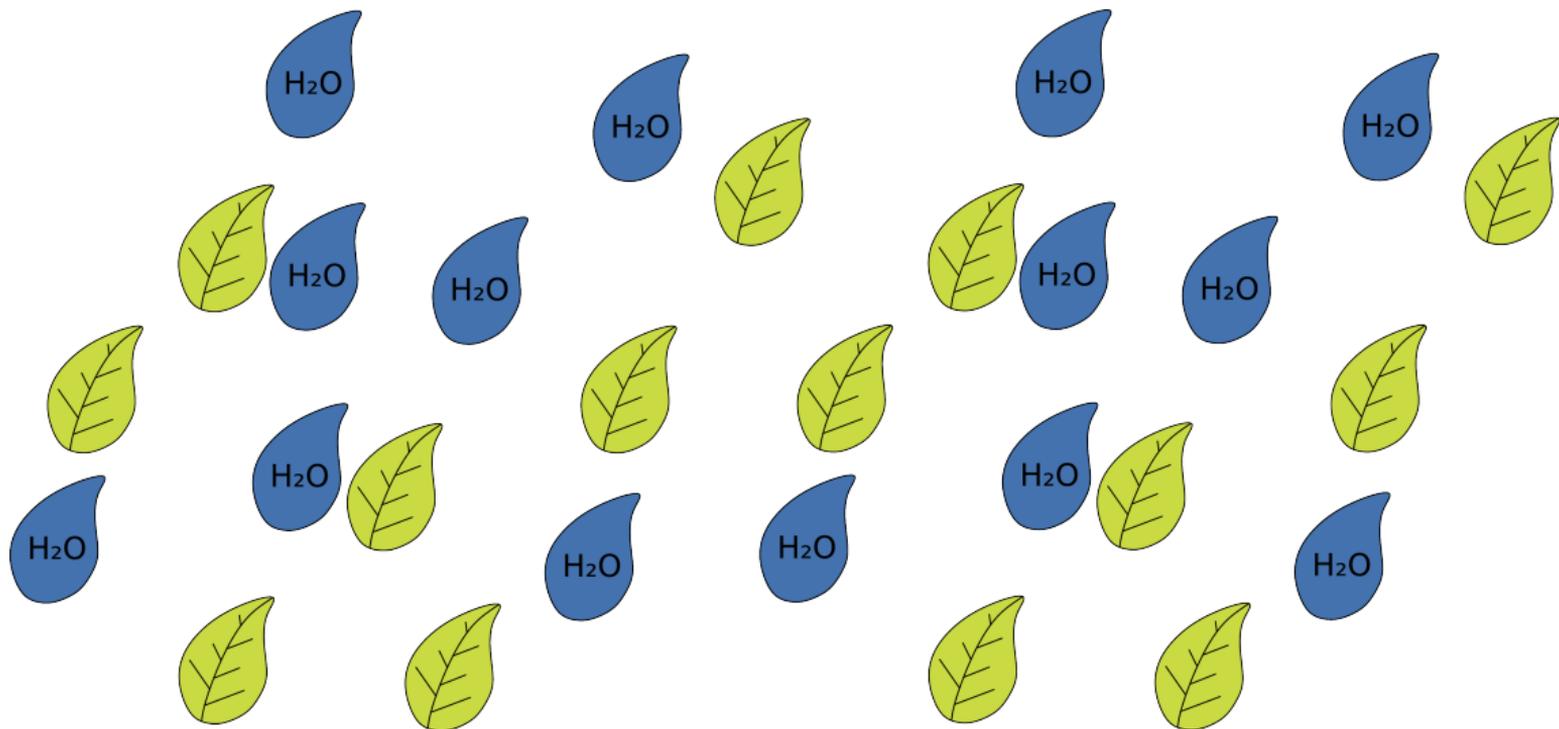


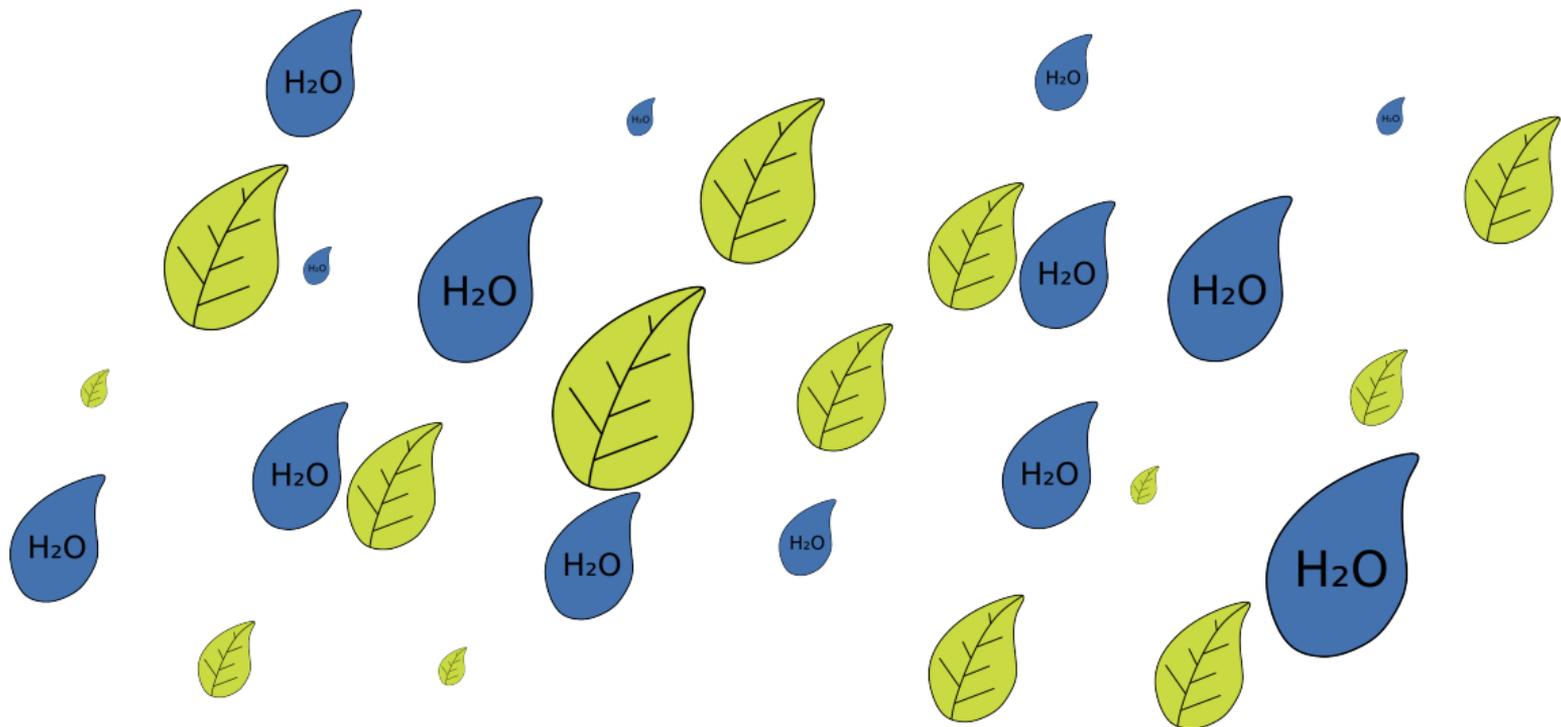
Images Courtesy of Florian Beier (TU Berlin)

Introductory Example: Water Your Plants Optimal Transport Style

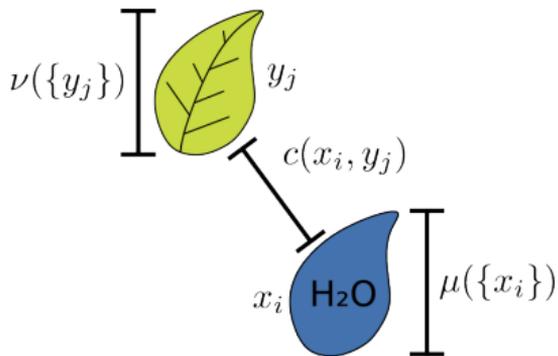


Images Courtesy of Florian Beier (TU Berlin)



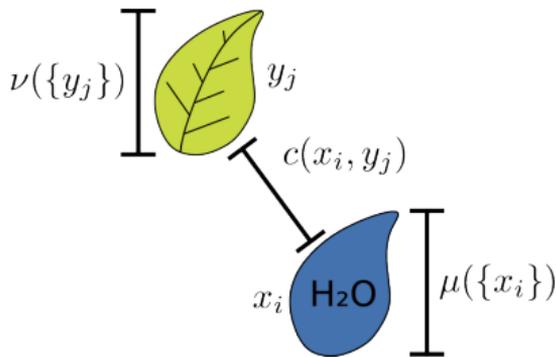


Kantorovich Problem



- positions water: $x_1, \dots, x_n \in \mathbb{R}^2$
- positions plants: $y_1, \dots, y_n \in \mathbb{R}^2$
- cost between x_i and y_j : $c(x_i, y_j) \in [0, \infty)$
- amount of water at x_i : $\mu(\{x_i\}) \in [0, 1]$
- demand of plant y_j : $\nu(\{y_j\}) \in [0, 1]$
- water moved from x_i to y_j : transport plan $\pi_{i,j}$

Kantorovich Problem



- positions water: $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^2$
- positions plants: $\mathbf{y}_1, \dots, \mathbf{y}_n \in \mathbb{R}^2$
- cost between \mathbf{x}_i and \mathbf{y}_j : $c(\mathbf{x}_i, \mathbf{y}_j) \in [0, \infty)$
- amount of water at \mathbf{x}_i : $\mu(\{\mathbf{x}_i\}) \in [0, 1]$
- demand of plant \mathbf{y}_j : $\nu(\{\mathbf{y}_j\}) \in [0, 1]$
- water moved from \mathbf{x}_i to \mathbf{y}_j : transport plan $\pi_{i,j}$

- Kantorovich Optimal Transport (OT) problem (1942):

$$\min_{\pi \in \mathbb{R}_{\geq 0}^{n \times m}} \left\{ \sum_{i=1}^n \sum_{j=1}^m \underbrace{c(\mathbf{x}_i, \mathbf{y}_j) \pi_{i,j}}_{\text{weighted transport-costs between } \mathbf{x}_i \text{ and } \mathbf{y}_i} : \sum_{j=1}^m \pi_{i,j} = \mu_i, \sum_{i=1}^n \pi_{i,j} = \nu_j \right\}$$

where $\mu_i = \mu(\{\mathbf{x}_i\})$, $\nu_j = \nu(\{\mathbf{y}_j\})$.

$$\pi \begin{array}{c} \begin{array}{ccc|c} 1/4 & 1/4 & 1/2 & \mu \\ \hline 1/2 & = & 1/8 + 1/4 + 1/8 & \\ 1/4 & + & 1/8 & 0 & 1/8 \\ 1/4 & + & 0 & 0 & 1/4 \\ \hline \nu & & & & \end{array} \end{array}$$

Kantorovich Problem (Formulation with Measures)

- $\mathcal{P}(\mathbb{X})$ probability measures on compact metric space \mathbb{X}
- $P_1(x, y) := x$ projection to the first component
- **Pushforward** of $f: \mathbb{X} \rightarrow \mathbb{Y}$ and $\mu \in \mathcal{P}(\mathbb{X})$ is

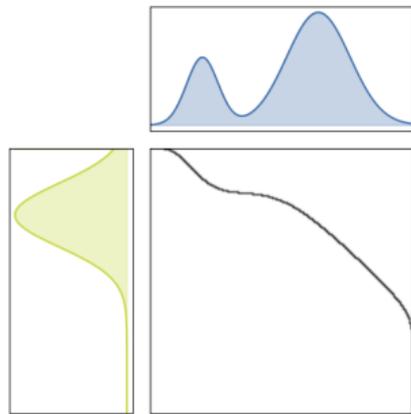
$$f_{\#}\mu := \mu \circ f^{-1} \in \mathcal{P}(\mathbb{Y})$$

- $\mu \in \mathcal{P}(\mathbb{X}), \nu \in \mathcal{P}(\mathbb{Y})$

$$\min_{\pi \in \mathcal{P}(\mathbb{X} \times \mathbb{Y})} \left\{ \int_{\mathbb{X} \times \mathbb{Y}} c(\mathbf{x}, \mathbf{y}) d\pi(\mathbf{x}, \mathbf{y}) : P_{1\#}\pi = \mu, P_{2\#}\pi = \nu \right\}$$

- If $\mathbb{X}, \mathbb{Y} \subset \mathbb{R}^d$ and $c(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2^p$, the minimum is the **Wasserstein- p distance**

$$W_p^p(\mu, \nu) := \min_{\pi \in \mathcal{P}(\mathbb{X} \times \mathbb{Y})} \left\{ \int_{\mathbb{X} \times \mathbb{Y}} \|\mathbf{x} - \mathbf{y}\|_2^p d\pi(\mathbf{x}, \mathbf{y}) : P_{1\#}\pi = \mu, P_{2\#}\pi = \nu \right\}$$



Wasserstein Barycenters

[Agueh Carlier 2011]

- $\mu_i \in \mathcal{P}(\mathbb{R}^d)$ and $t_i \in [0, 1]$, $i = 1, \dots, N$ with $\sum_{i=1}^N t_i = 1$
- Wasserstein Barycenter

$$\operatorname{argmin}_{\nu \in \mathcal{P}(\mathbb{R}^d)} \sum_{i=1}^N t_i W_2^2(\mu_i, \nu)$$

- For $N = 2$ inputs, a barycenter $\hat{\nu}$ can be computed by

$$\hat{\nu} = \sum_{(\mathbf{x}_1, \mathbf{x}_2) \in \operatorname{supp}(\hat{\pi})} \hat{\pi}(\{(\mathbf{x}_1, \mathbf{x}_2)\}) \delta_{\{t_1 \mathbf{x}_1 + t_2 \mathbf{x}_2\}}, \quad (\text{McCann Interpolation})$$

where $\hat{\pi}$ solves $W_2(\mu_1, \mu_2)$

- Wasserstein distances and barycenters on \mathbb{R}^d are computationally expensive



Wasserstein barycenters
between 3 measures

[Bonneel Rabin Peyré Pfister 2015]

Content

- 1 Introduction to Optimal Transport
- 2 Sliced and 1-dimensional Optimal Transport**
- 3 Sliced Optimal Transport on the Sphere
- 4 Numerics

Sliced Optimal Transport on \mathbb{R}^d

[Rabin Peyré Delon Bernot 2012] [Kolouri Park Rohde 2016]

- **Absolutely continuous** measure $\mu \in \mathcal{P}_{\text{ac}}(\mathbb{R}^d)$ with density f_μ , i.e.

$$\mu(A) = \int_A f_\mu(x) \, dx \quad \forall A \in \mathcal{B}(\mathbb{R}^d)$$

- **Radon transform**

$$\mathcal{R}_\theta f(t) := \int_{\theta^\top \mathbf{x} = t} f(\mathbf{x}) \, d\mathbf{x}, \quad \theta \in \mathbb{S}^{d-1}, t \in \mathbb{R}$$

- **Sliced Wasserstein distance**

$$\text{SW}_p^p(\mu, \nu) := \int_{\mathbb{S}^{d-1}} W_p^p(\mathcal{R}_\theta f_\mu, \mathcal{R}_\theta f_\nu) \, d\xi$$

- Integral over Wasserstein distances of one-dimensional measures (density functions)

OT on \mathbb{R}

[Vilani 03] [Peyre Cuturi 19]

- Absolutely continuous measure $\mu \in \mathcal{P}_{\text{ac}}(\mathbb{R})$
- Cumulative distribution function $F_\mu(x) := \mu((-\infty, x])$
- Quantile function $F_\mu^{-1}(r) := \min\{x \in \mathbb{R} : F_\mu(x) \geq r\}$, $r \in [0, 1]$
- Wasserstein distance

$$W_p^p(\mu, \nu) = \int_0^1 |\tilde{F}_\mu^{-1}(r) - \tilde{F}_\nu^{-1}(r)|^p dr$$

- Unique OT plan

$$\pi = (\text{Id}, T^{\mu, \nu})_{\#} \mu \quad \text{with} \quad T^{\mu, \nu}(x) := F_\nu^{-1}(F_\mu(x)), \quad x \in \mathbb{R}$$

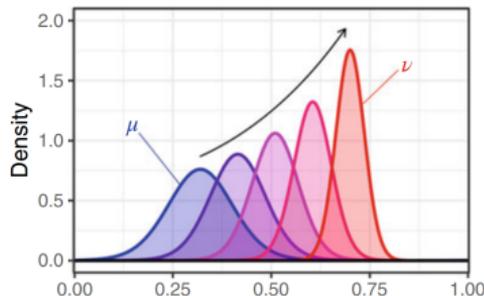
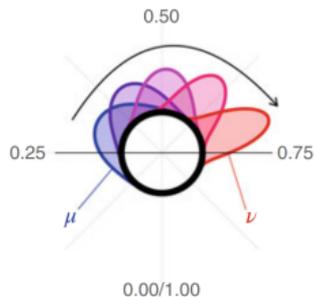
OT on the Circle

[Delon Salomon Sobolevski 2010] [Rabin Delon Gousseau 2011]

- Circle $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$
- Idea: Cut \mathbb{T} at the right position to get an OT problem on the interval
- Wasserstein distance

$$W_p^p(\mu, \nu) = \min_{\theta \in \mathbb{R}} \int_0^1 |\tilde{F}_\mu^{-1}(r) - (\tilde{F}_\nu - \theta)^{-1}(r)|^p dr$$

- Compute optimal θ via bisection



Barycenters of measures μ, ν on \mathbb{T} , see [Hundrieser Klatt Munk 2022]



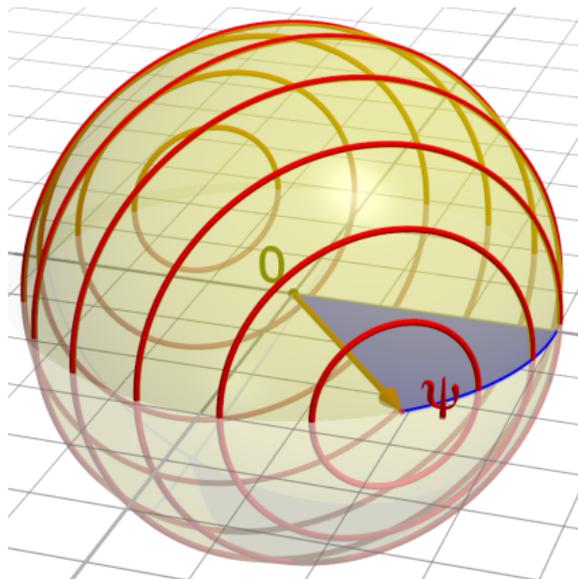
Content

- 1 Introduction to Optimal Transport
- 2 Sliced and 1-dimensional Optimal Transport
- 3 Sliced Optimal Transport on the Sphere**
- 4 Numerics

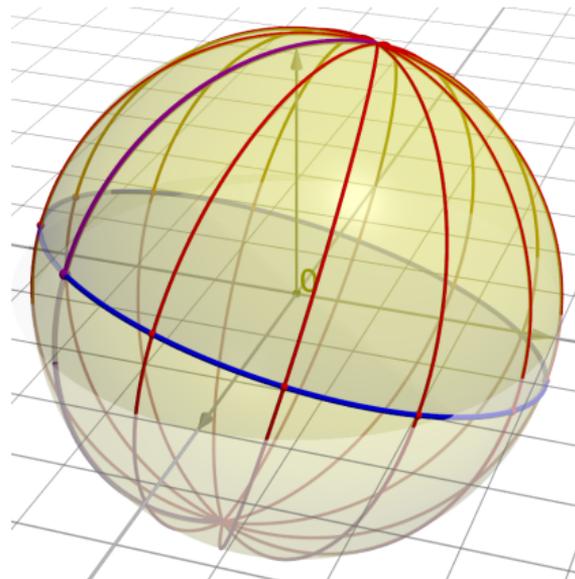


Sliced OT on the Sphere $\mathbb{S}^2 := \{\xi \in \mathbb{R}^3 : \|\xi\| = 1\}$

Vertical Slices (“egg cutter”)



Semicircles (“orange segments”)



Vertical Slices

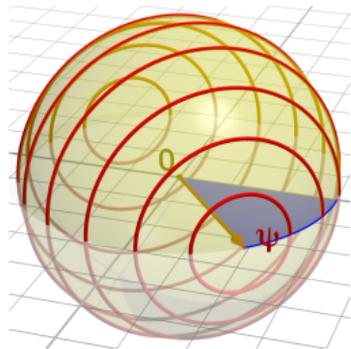
- Slicing operator for fixed $\psi \in \mathbb{T} \sim [0, 2\pi)$

$$\mathcal{S}_\psi : \mathbb{S}^2 \rightarrow [-1, 1], \quad \mathcal{S}_\psi(\boldsymbol{\xi}) := \boldsymbol{\xi}^\top (\cos \psi, \sin \psi, 0),$$

- Circle

$$\mathcal{S}_\psi^{-1}(t) = \{\boldsymbol{\xi} \in \mathbb{S}^2 : \mathcal{S}_\psi(\boldsymbol{\xi}) = t\}, \quad t \in [-1, 1],$$

is the intersection of \mathbb{S}^2 and the plane with normal ψ and distance t from the origin.



- Vertical slice transform

[Gindikin Reeds Shepp 1994] [Zangerl Scherzer 2010]

$$\mathcal{V}f(\psi, t) := \frac{1}{|\mathbb{S}^2| \sqrt{1-t^2}} \int_{\mathcal{S}_\psi^{-1}(t)} f(\boldsymbol{\xi}) \, ds(\boldsymbol{\xi}), \quad \psi \in \mathbb{T}, \quad t \in (-1, 1)$$

- Restriction $\mathcal{V}_\psi := 2\pi \mathcal{V}(\psi, \cdot)$ transforms a function on \mathbb{S}^2 to many functions on $[-1, 1]$
- Singular value decomposition of \mathcal{V} is known

[Hielscher Q. 2015]

Vertical Slices (Definition for Measures)

We generalize the **vertical slice transform** \mathcal{V}_ψ for measures by

$$\begin{aligned} \mathcal{V} : \mathcal{M}(\mathbb{S}^2) &\rightarrow \mathcal{M}(\mathbb{T} \times [-1, 1]), & \mu &\mapsto T_{\#}(u_{\mathbb{T}} \times \mu) \quad \text{with} \quad T(\psi, \xi) := (\psi, \mathcal{S}_\psi(\xi)), \\ \mathcal{V}_\psi : \mathcal{M}(\mathbb{S}^2) &\rightarrow \mathcal{M}([-1, 1]), & \mu &\mapsto (\mathcal{S}_\psi)_{\#}\mu = \mu \circ \mathcal{S}_\psi^{-1}. \end{aligned}$$

where $u_{\mathbb{T}}$ is the normalized Lebesgue measure on \mathbb{T} .

Theorem (absolutely continuous measures)

For $f \in L^1(\mathbb{S}^2)$,

$$\mathcal{V}[f\sigma_{\mathbb{S}^2}] = (\mathcal{V}f)\sigma_{\mathbb{T} \times [-1, 1]},$$

where σ is the Lebesgue measure.

Theorem

The vertical slice transform $\mathcal{V} : \mathcal{M}_{\text{sym}}(\mathbb{S}^2) \rightarrow \mathcal{M}(\mathbb{T} \times [-1, 1])$ is injective, where

$$\mathcal{M}_{\text{sym}}(\mathbb{S}^2) := \{\mu \in \mathcal{M}(\mathbb{S}^2) : \langle \mu, f \rangle = \langle \mu, \check{f} \rangle \forall f \in C(\mathbb{S}^2)\}, \quad \check{f}(\mathbf{x}) = f(x_1, x_2, -x_3),$$

Semicircle transform

- Spherical coordinates $\Phi: \mathbb{T} \times [0, \pi] \rightarrow \mathbb{S}^2$,

$$\Phi(\varphi, \vartheta) := (\cos \varphi \sin \vartheta, \sin \varphi \sin \vartheta, \cos \vartheta) \in \mathbb{S}^2$$

- Rotation group

$$\text{SO}(3) := \{Q \in \mathbb{R}^{3 \times 3} : Q^T Q = I, \det Q = 1\}$$

- Euler angles

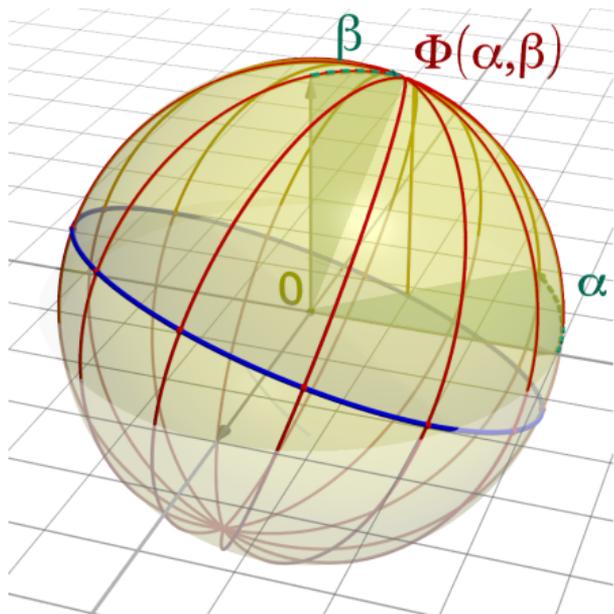
$$\Psi(\alpha, \beta, \gamma) := R_3(\alpha)R_2(\beta)R_3(\gamma) \in \text{SO}(3),$$

where $R_i(\alpha)$ is the rotation around the i -th axis with angle α

- Normalized semicircle transform \mathcal{W} of $f: \mathbb{S}^2 \rightarrow \mathbb{R}$

$$\mathcal{W}f(\alpha, \beta, \gamma) := \frac{1}{4\pi} \int_0^\pi f(\Psi(\alpha, \beta, \vartheta) \Phi(\gamma, \vartheta)) \sin \vartheta \, d\vartheta$$

- Restriction $\mathcal{W}_{\alpha, \beta} f := 4\pi \mathcal{W}f(\alpha, \beta, \cdot)$
- Unnormalized semicircle transform (without $\sin \vartheta$) due to [\[Groemer 1998\]](#)



Theorem (singular value decomposition)

[Q. Beinert Steidl 2023]

The normalized semicircle transform fulfills

$$\mathcal{W}Y_n^k = w_n Z_n^k, \quad n \in \mathbb{N}_0, \quad k \in \{-n, \dots, n\},$$

with the singular values $w_n := \|\mathcal{W}Y_n^k\|_{L^2(\mathbb{S}^2)} \in \mathcal{O}(n^{-1/2})$ and the orthonormal functions

$$Z_n^k := w_n^{-1} \sum_{j=-n}^n \lambda_n^j \overline{D_n^{k,j}} \in L^2(\text{SO}(3)),$$

where $\lambda_0^0 := 2(4\pi)^{-3/2}$ and, for $n \in \mathbb{N}$ and $j \in \{1, \dots, n\}$ with $n+j$ even,

$$\lambda_n^j := \frac{(-1)^j}{4\pi} \sqrt{\frac{2n+1}{4\pi} \frac{(n-j)!}{(n+j)!} \frac{j(n-2)!!(n+j-1)!!}{(n-j)!!(n+1)!!}} \begin{cases} 2 & : \quad n \text{ even,} \\ \pi & : \quad n \text{ odd,} \end{cases}$$

$\lambda_n^{-j} := (-1)^j \lambda_n^j$, and $\lambda_n^j = 0$ otherwise. Here Y_n^k denote the spherical harmonics and $D_n^{k,j}$ the rotational harmonics (Wigner D-functions). Moreover, there are constants $C_1, C_2 > 0$ such that

$$C_1 (n+1)^{-1/2} \leq w_n \leq C_2 (n+1)^{-1/2} \quad \forall n \in \mathbb{N}_0.$$

Semicircle Transform of Measures

For $\Phi(\alpha, \beta) \in \mathbb{S}^2$, we define the **azimuth operator**

$$\mathcal{A}_{\alpha, \beta}: \mathbb{S}^2 \rightarrow \mathbb{T}, \quad \mathcal{A}_{\alpha, \beta}(\boldsymbol{\xi}) := \text{azi}(\Psi(\alpha, \beta, 0)^T \boldsymbol{\xi}),$$

where azi denotes the first component of the inverse of the spherical coordinate transform Φ , i.e.

$$\text{azi}(\Phi(\varphi, \theta)) = \varphi \quad \forall \varphi, \theta.$$

We generalize the (restricted) **semicircle transform** $\mathcal{W}_{\alpha, \beta}$ by

$$\begin{aligned} \mathcal{W}_{\alpha, \beta}: \mathcal{M}(\mathbb{S}^2) &\rightarrow \mathcal{M}(\mathbb{T}), & \mu &\mapsto (\mathcal{A}_{\alpha, \beta})_{\#} \mu = \mu \circ \mathcal{A}_{\alpha, \beta}^{-1}. \\ \mathcal{W}: \mathcal{M}(\mathbb{S}^2) &\rightarrow \mathcal{M}(\text{SO}(3)), & \mu &\mapsto (T_{\mathcal{W}})_{\#} (u_{\mathbb{S}^2} \times \mu) \quad \text{with} \quad T_{\mathcal{W}}(\Phi(\alpha, \beta), \boldsymbol{\xi}) := \Psi(\alpha, \beta, \mathcal{A}_{\alpha, \beta}(\boldsymbol{\xi})). \end{aligned}$$

Theorem (injectivity)

The normalized semicircle transform is injective $\mathcal{M}(\mathbb{S}^2) \rightarrow \mathcal{M}(\text{SO}(3))$.

Sliced Spherical Wasserstein Distances

Definition

Let $p \in [1, \infty)$ and $\mu, \nu \in \mathcal{M}(\mathbb{S}^2)$. We define the **vertical sliced Wasserstein distance**

$$\text{VSW}_p^p(\mu, \nu) := \frac{1}{2\pi} \int_0^{2\pi} W_p^p(\mathcal{V}_\psi \mu, \mathcal{V}_\psi \nu) \, d\psi$$

and the **semicircular sliced Wasserstein distance**

$$\text{SSW}_p^p(\mu, \nu) := \int_{\mathbb{S}^2} W_p^p(\mathcal{W}_{\alpha, \beta} \mu, \mathcal{W}_{\alpha, \beta} \nu) \, du_{\mathbb{S}^2}(\Phi(\alpha, \beta)),$$

which are integrals over Wasserstein distances on $[-1, 1]$ and \mathbb{T} , respectively.

SSW was defined in [\[Bonet Berg Courty Septier Drumetz Pham 2023\]](#)

Theorem (metric properties)

[Q. Beinert Steidl 2023]

Let $p \in [1, \infty)$. The vertical sliced Wasserstein distance VSW_p is a metric on $\mathcal{M}_{\text{sym}}(\mathbb{S}^2)$, and the semicircular Wasserstein distance SSW_p is a metric on $\mathcal{M}(\mathbb{S}^2)$.

Invariance to Rotations

Theorem

Let $p \in [1, \infty)$. The vertical sliced Wasserstein distance VSW_p is invariant to rotations \mathbf{R}_3 around the vertical axis, i.e., for all $\mu, \nu \in \mathcal{M}(\mathbb{S}^2)$ and $\alpha \in \mathbb{T}$,

$$VSW_p(\mu, \nu) = VSW_p(\mu \circ \mathbf{R}_3(\alpha), \nu \circ \mathbf{R}_3(\alpha)).$$

Theorem

Let $p \in [1, \infty)$. The semicircular sliced Wasserstein distance SSW_p is rotationally invariant, i.e., for all $\mu, \nu \in \mathcal{M}(\mathbb{S}^2)$ and $\mathbf{Q} \in \text{SO}(3)$,

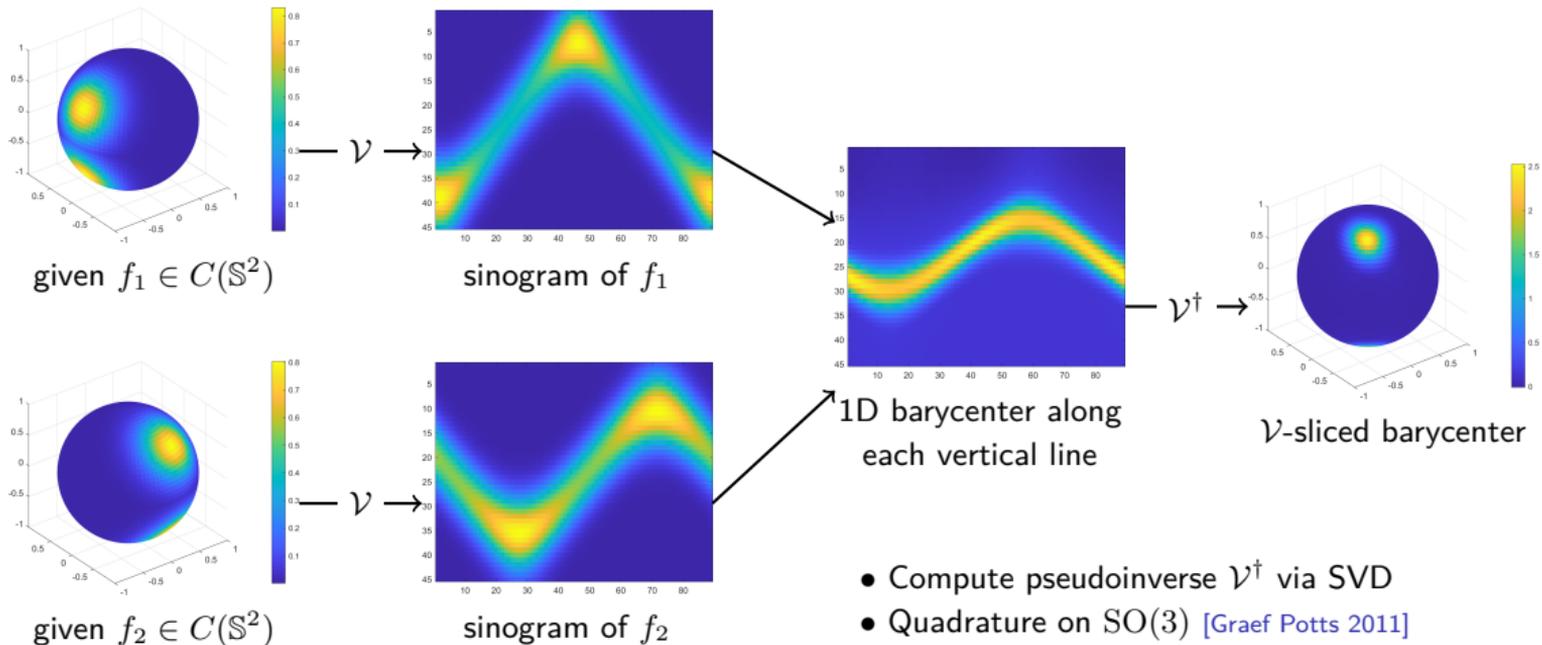
$$SSW_p(\mu, \nu) = SSW_p(\mu \circ \mathbf{Q}, \nu \circ \mathbf{Q}).$$



Content

- 1 Introduction to Optimal Transport
- 2 Sliced and 1-dimensional Optimal Transport
- 3 Sliced Optimal Transport on the Sphere
- 4 Numerics**

Sliced (Radon) barycenter



Inversion of \mathcal{W} by Variational Approach

- Barycenter of probability measures is again probability measure
- $\mathcal{W}^\dagger g$ might be negative even if $g \geq 0$, therefore not a probability density
- Find approximate solution f of the inversion problem $\mathcal{W}f = g$ via

$$\operatorname{argmin}_{\substack{f \geq 0 \\ \int_{\mathbb{S}^2} f = 1}} \operatorname{KL}(\mathcal{W}f, g) + \rho \operatorname{KL}(f, 1),$$

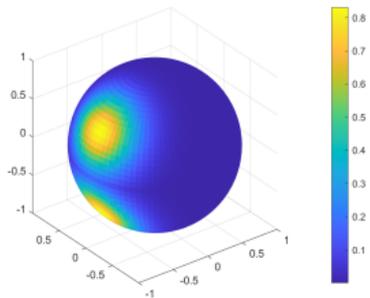
with the **Kullback–Leibler (KL) divergence**

$$\operatorname{KL}(f, \tilde{f}) := \langle f, \log f - \log \tilde{f} \rangle + \langle \tilde{f} - f, 1 \rangle$$

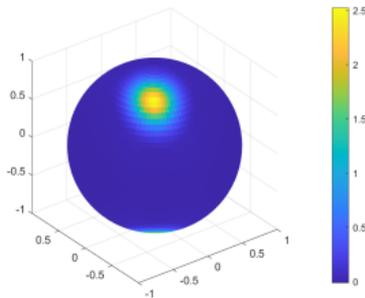
for $f, \tilde{f} \geq 0$ with $f(x) = 0$ whenever $\tilde{f}(x) = 0$ and $0 \log 0 := 0$.

- Primal-dual splitting [[Chambolle Pock 2016](#)] yields converging iteration

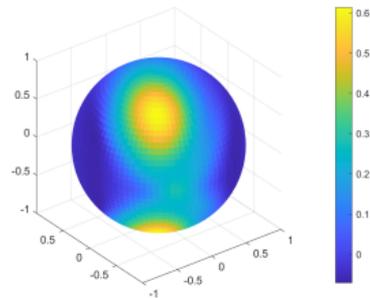
$$\begin{aligned} \mathbf{f}^{k+1} &:= \operatorname{proj}_{\Delta}(\mathbf{f}^k - \tau \mathcal{W}^* \mathbf{y}_1^k - \tau \mathbf{y}_2^k), & \tilde{\mathbf{f}}^{k+1} &:= \mathbf{f}^{k+1} + \theta(\mathbf{f}^{k+1} - \mathbf{f}^k), \\ \mathbf{y}_1^{k+1} &:= \operatorname{prox}_{\sigma \operatorname{KL}^*(\cdot, g)}(\mathbf{y}_1^k + \sigma \mathcal{W} \tilde{\mathbf{f}}^{k+1}), & \mathbf{y}_2^{k+1} &:= \operatorname{prox}_{\sigma(\rho \operatorname{KL})^*(\cdot, 1)}(\mathbf{y}_2^k + \sigma \tilde{\mathbf{f}}^{k+1}). \end{aligned}$$



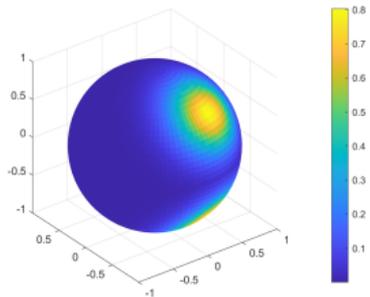
(a) Given measure μ
(van Mises–Fischer distribution)



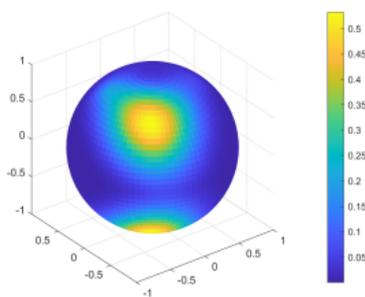
(b) \mathcal{V} -sliced barycenter (0.01 s)



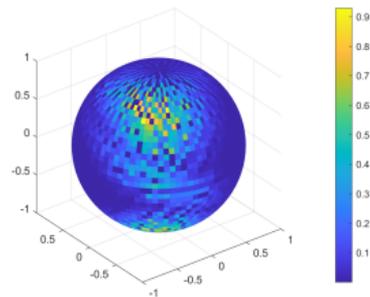
(c) \mathcal{W} -sliced barycenter (3.2 s)



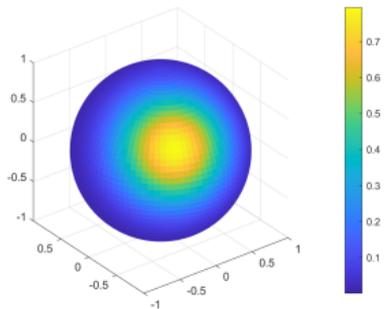
(d) Given measure ν



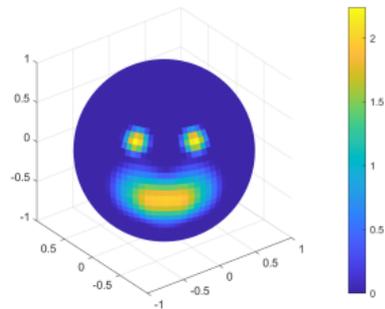
(e) Regularized Wasserstein
barycenter (PythonOT, 19 s)



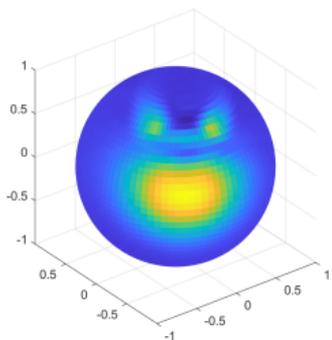
(f) Unregularized Wasserstein
barycenter (PythonOT, 33 h)



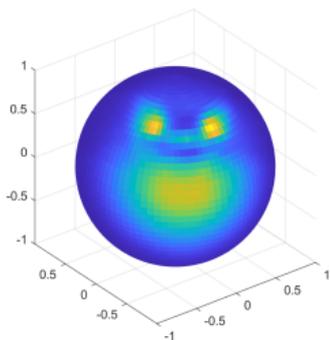
(a) Density of vMF distribution μ



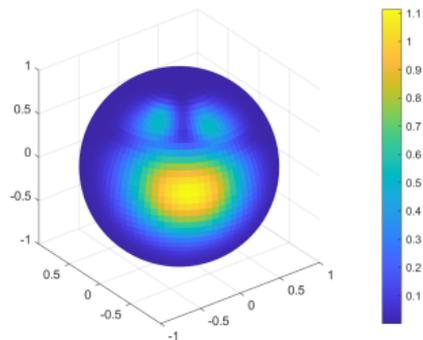
(b) Density of ν (quadratic spline)



(c) \mathcal{W} -sliced barycenter with Moore–Penrose inverse



(d) \mathcal{W} -sliced barycenter with regularized inverse



(e) Regularized Wasserstein barycenter



Conclusions

- We defined the vertical slice and normalized semicircle transform, and generalized them to measures.
- For absolutely continuous measures, the generalized and initial definitions coincide
- We showed an SVD of the normalized semicircle transform, which provides an approach for numerical computations and inversion.
- The normalized semicircle transform is injective for measures, hence the sliced Wasserstein distance fulfills the properties of a metric.



M Quellmalz, R Beinert, G Steidl. Sliced optimal transport on the sphere. *Inverse Problems* 39, 2023.

Future

- Modify vertical slice transform to support any direction (not just on the equator)
- Other manifolds, \mathbb{S}^d or $SO(3)$

Conclusions

- We defined the vertical slice and normalized semicircle transform, and generalized them to measures.
- For absolutely continuous measures, the generalized and initial definitions coincide
- We showed an SVD of the normalized semicircle transform, which provides an approach for numerical computations and inversion.
- The normalized semicircle transform is injective for measures, hence the sliced Wasserstein distance fulfills the properties of a metric.



M Quellmalz, R Beinert, G Steidl. Sliced optimal transport on the sphere. *Inverse Problems* 39, 2023.

Future

- Modify vertical slice transform to support any direction (not just on the equator)
- Other manifolds, \mathbb{S}^d or $SO(3)$

Thank you for your attention!