Fast Optimal Transport through Sliced Generalized Wasserstein Geodesics

Joint work with Guillaume Mahey, Gilles Gasso, Clément Bonet and Nicolas Courty
NeurIPS 2023 [5]

Laetitia Chapel
laetitia.chapel@irisa.fr

IRISA, Rennes, France

Institut Agro Rennes-Angers

Workshop on Optimal Transport: from theory to applications, Berlin 2024

Table of Contents

Background on Optimal Transport
Optimal transport and Wasserstein distance

- Optimal transport and Wasserstein distance

Linear loss

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma(\mu_1, \mu_2)} \int_{X \times Y} c(x, y) \, d\gamma(x, y)$$

where $\Gamma(\mu_1, \mu_2) \stackrel{\text{def}}{=} \{\gamma \in \mathcal{M}_+(X \times Y) \text{ s.t. } (\pi_x)_\# \gamma = \mu_1 \text{ and } (\pi_y)_\# \gamma = \mu_2\}$ with $\pi_x : X \times Y \to X$.

Marginal constraints

Background on Optimal Transport
Optimal transport and Wasserstein distance

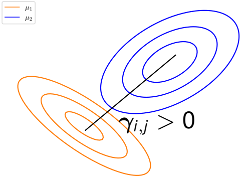
- Optimal transport and Wasserstein distance

Linear loss

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma(\mu_1,\mu_2)} \int_{X \times Y} c(x,y) \, d\gamma(x,y)$$

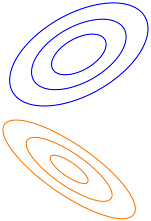
where $\Gamma(\mu_1, \mu_2) \stackrel{\text{def}}{=} \{\gamma \in \mathcal{M}_+(X \times Y) \text{ s.t. } (\pi_x)_\# \gamma = \mu_1 \text{ and } (\pi_y)_\# \gamma = \mu_2\}$ with $\pi_x : X \times Y \to X$.

Marginal constraints



with $(\pi_x)_\# \gamma = \mu_1$

and $(\pi_y)_\# \gamma = \mu_2$



The **transport plan** $\gamma(x,y)$ specifies for each pair (x,y) how many particles go from x to y

- Wasserstein distance when $c(x,y) = |x-y|^p$

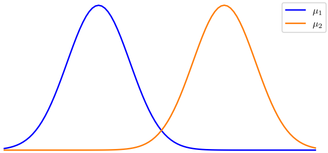
$$\mathcal{W}_p(\mu_1, \mu_2) \triangleq \left(\inf_{\gamma \in \Gamma(\mu_1,\mu_2)} \int_{X \times Y} c(x,y) \, d\gamma(x,y) \right)^{1/p}$$

Background on Optimal Transport
Transport map and Wasserstein Geodesics

- In some cases, the optimal plan $\gamma*$ is a Monge map of the form $(Id, \mathbf{T})\#\mu_1$, e.g. for $p = 2$

$$\mathcal{W}_p^p(\mu_1, \mu_2) \triangleq \inf_{\mathbf{T}} \int \|x - \mathbf{T}(x)\|_2^2 \, d\mu_1(x)$$

where **T** is a **transport map** and $\mathbf{T}_{\#}\mu_1 = \mu_2$

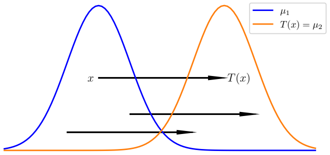


Background on Optimal Transport
Transport map and Wasserstein Geodesics

- In some cases, the optimal plan $\gamma*$ is a Monge map of the form $(Id, \mathbf{T})\#\mu_1$, e.g. for $p = 2$

$$\mathcal{W}_p^p(\mu_1, \mu_2) \triangleq \inf_{\mathbf{T}} \int \|x - \mathbf{T}(x)\|_2^2 \, d\mu_1(x)$$

where **T** is a **transport map** and $\mathbf{T}_{\#}\mu_1 = \mu_2$



Defines for each particle located at x what is its destination $\mathbf{T}(x)$

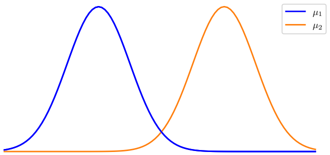
Background on Optimal Transport
Transport map and Wasserstein Geodesics

- In some cases, the optimal plan $\gamma*$ is a Monge map of the form $(Id, \mathbf{T})\#\mu_1$, e.g. for $p = 2$

$$\mathcal{W}_p^p(\mu_1, \mu_2) \triangleq \inf_{\mathbf{T}} \int \|x - \mathbf{T}(x)\|_2^2 \, d\mu_1(x)$$

where $\mathbf{T}$ is a **transport map** and $\mathbf{T}_{\#}\mu_1 = \mu_2$

- **Wasserstein geodesics** $\mu^{1\to 2}(t) \triangleq (t\mathbf{T}^{1\to 2} + (1-t)Id)\#\mu_1$ with $\mathbf{T}^{1\to 2}$ the optimal map



For short, we denote $\mu^{1\to 2}$ for $t = 0.5$

Background on Optimal Transport
Curvature of the Wasserstein space

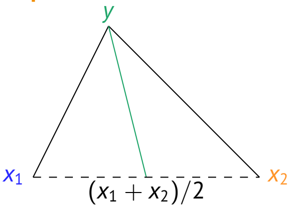
- The Wasserstein space is of positive curvature

$$\mathcal{W}_2^2(\mu^{1\to 2}, \nu) \geq \frac{1}{2}\mathcal{W}_2^2(\mu_1, \nu) + \frac{1}{2}\mathcal{W}_2^2(\nu, \mu_2) - \frac{1}{4}\mathcal{W}_2^2(\mu_1, \mu_2)$$

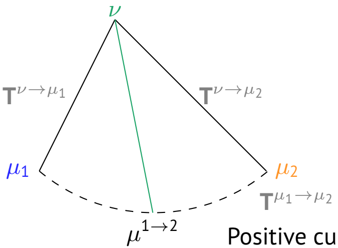
or equivalently

$$\mathcal{W}_2^2(\mu_1, \mu_2) \geq 2\mathcal{W}_2^2(\mu_1, \nu) + 2\mathcal{W}_2^2(\nu, \mu_2) - 4\,\mathcal{W}_2^2(\mu^{1\to 2}, \nu)$$

for ν a **pivot measure**.



Parallelogram law in $\mathbb{R}^d$



Positive curvature of $\mathcal{W}$ space

Background on Optimal Transport
Curvature of the Wasserstein space

- The Wasserstein space is of positive curvature

$$\mathcal{W}_2^2(\mu^{1\to 2},\nu) \geq \frac{1}{2}\mathcal{W}_2^2(\mu_1,\nu) + \frac{1}{2}\mathcal{W}_2^2(\nu,\mu_2) - \frac{1}{4}\mathcal{W}_2^2(\mu_1,\mu_2)$$

or equivalently

$$\mathcal{W}_2^2(\mu_1,\mu_2) \geq 2\mathcal{W}_2^2(\mu_1,\nu) + 2\mathcal{W}_2^2(\nu,\mu_2) - 4\,\mathcal{W}_2^2(\mu^{1\to 2},\nu)$$

for ν a **pivot measure**.

- The Wasserstein space is flat when μ_1,μ_2,ν are 1d

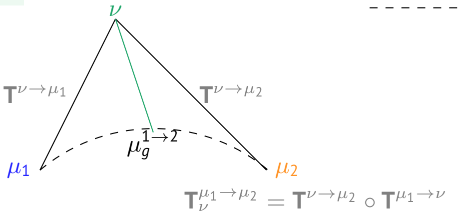
$$\mathcal{W}_2^2(\mu_1,\mu_2) = 2\mathcal{W}_2^2(\mu_1,\nu) + 2\mathcal{W}_2^2(\nu,\mu_2) - 4\mathcal{W}_2^2(\mu^{1\to 2},\nu)$$

Background on Optimal Transport
Wasserstein Generalized Geodesics

- Has been introduced by Ambrosio et al. [1]
- Wasserstein Geodesic: $\mu^{1\to 2}(t) \triangleq (t\ \mathbf{T}^{1\to 2} + (1-t)Id)\#\mu_1$
- **Wasserstein Generalized Geodesic**: $\mu_g^{1\to 2}(t) \triangleq (t\ \mathbf{T}^{\nu\to\mu_2} + (1-t)\ \mathbf{T}^{\nu\to\mu_1})\#\nu$ for ν a **pivot measure**.

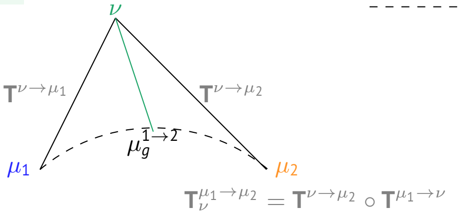
Background on Optimal Transport
Wasserstein Generalized Geodesics

- Has been introduced by Ambrosio et al. [1]
- Wasserstein Geodesic: $\mu^{1\to 2}(t) \triangleq (t\,\mathbf{T}^{1\to 2} + (1-t)Id)\#\mu_1$
- **Wasserstein Generalized Geodesic**: $\mu_g^{1\to 2}(t) \triangleq (t\,\mathbf{T}^{\nu\to\mu_2} + (1-t)\,\mathbf{T}^{\nu\to\mu_1})\#\nu$ for ν a **pivot measure**.
- Negative curvature $\mathcal{W}_2^2(\mu_g^{1\to 2},\nu) \leq \frac{1}{2}\mathcal{W}_2^2(\mu_1,\nu) + \frac{1}{2}\mathcal{W}_2^2(\nu,\mu_2) - \frac{1}{4}\mathcal{W}_2^2(\mu_1,\mu_2)$



Background on Optimal Transport
Wasserstein Generalized Geodesics

- Has been introduced by Ambrosio et al. [1]
- Wasserstein Geodesic: $\mu^{1\to 2}(t) \triangleq (t\, \mathbf{T}^{1\to 2} + (1-t)Id)\#\mu_1$
- **Wasserstein Generalized Geodesic**: $\mu_g^{1\to 2}(t) \triangleq (t\, \mathbf{T}^{\nu\to\mu_2} + (1-t)\, \mathbf{T}^{\nu\to\mu_1})\#\nu$ for ν a **pivot measure**.
- Negative curvature $\mathcal{W}_2^2(\mu_g^{1\to 2}, \nu) \leq \frac{1}{2}\mathcal{W}_2^2(\mu_1, \nu) + \frac{1}{2}\mathcal{W}_2^2(\nu, \mu_2) - \frac{1}{4}\mathcal{W}_2^2(\mu_1, \mu_2)$



- $\nu-$**Wasserstein distance**: $\mathcal{W}_\nu^2(\mu_1, \mu_2) = 2\mathcal{W}_2^2(\mu_1, \nu) + 2\mathcal{W}_2^2(\nu, \mu_2) - 4\mathcal{W}_2^2(\mu_g^{1\to 2}, \nu)$ with $\mathcal{W}_\nu^2(\mu_1, \mu_2) \geq \mathcal{W}_2^2(\mu_1, \mu_2)$

Computational Optimal Transport
Discrete formulation of OT

- For $\mu_1 = \sum_{i=1}^{n} h_i \delta_{x_i}$ and $\mu_2 = \sum_{j=1}^{m} g_j \delta_{y_j}$ and a quadratic cost, we solve

$$\mathcal{W}_2^2(\mu_1, \mu_2) \triangleq \min_{\gamma \in \Gamma(\mu_1, \mu_2)} \sum_{i,j} c(x_i, y_j) \gamma_{i,j}$$

$\rightarrow$ linear solvers with $O(n^3 \log(n))$ complexity

Computational Optimal Transport

Discrete formulation of OT

- For $\mu_1 = \sum_{i=1}^{n} h_i \delta_{x_i}$ and $\mu_2 = \sum_{j=1}^{m} g_j \delta_{y_j}$ and a quadratic cost, we solve

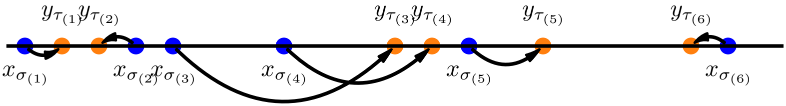
$$\mathcal{W}_2^2(\mu_1, \mu_2) \triangleq \min_{\gamma \in \Gamma(\mu_1, \mu_2)} \sum_{i,j} c(x_i, y_j)\gamma_{i,j}$$

  $\rightarrow$ linear solvers with $O(n^3 \log(n))$ complexity

- When μ_1 and μ_2 are 1D distributions and $n = m$ with uniform masses, the solution is given by

$$\mathcal{W}_2^2(\mu_1, \mu_2) \triangleq \frac{1}{n} \sum_{i=1}^{n} (x_{\sigma(i)} - y_{\tau(i)})^2$$

  $\rightarrow$ the optimal transport plan respects the ordering of the elements $x_{\sigma(i-1)} \leq x_{\sigma(i)}$ and $y_{\tau(i-1)} \leq y_{\tau(i)}$, complexity $O(n \log(n))$ and $O(n + n \log(n))$ for computing the distance



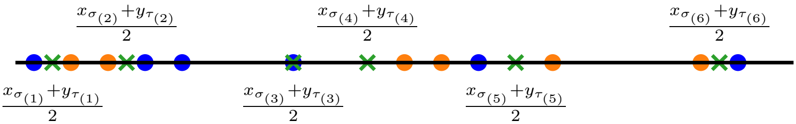
Computational Optimal Transport
Geodesic in 1D

- In 1D, the middle of the geodesic can be easily computed

$$(x_{\sigma(i)} + y_{\tau(i)})/2$$

- And when we take the pivot measure ν to be the middle of the geodesic $\mu^{1\to 2}$, we have

$$\mathcal{W}_2^2(\mu_1, \mu_2) = \mathcal{W}_\nu^2(\mu_1, \mu_2) = 2\mathcal{W}_2^2(\mu_1, \nu) + 2\mathcal{W}_2^2(\nu, \mu_2)$$

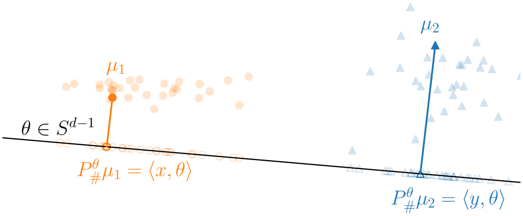


Computational Optimal Transport
Sliced Wasserstein on $\mathbb{R}^d$

1. Slice the distribution along lines $\theta \in S^{d-1}$
2. Project μ_1 and μ_2 onto θ: $P^\theta_\# \mu$, with $P^\theta : \mathbb{R}^d \to \mathbb{R}, \boldsymbol{x} \mapsto \langle \boldsymbol{x}, \theta \rangle$
3. Compute 1d Wasserstein onto the projected samples in 1d
4. Average all the distances

$$\mathcal{SW}_2^2(\mu_1, \mu_2) \triangleq \int_{S^{d-1}} W_2^2(P^\theta_\# \mu_1, P^\theta_\# \mu_2) d\omega(\theta),$$

with ω uniform distribution on S^{d-1}.



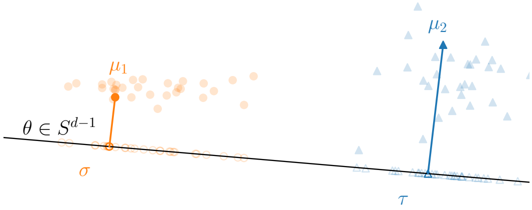
$\to$ provides a lower bound of $W_2^2(\mu_1, \mu_2)$ with complexity $O(Ln + Ln \log(n))$, L number of lines

Computational Optimal Transport
Projected Wasserstein Distance on $\mathbb{R}^d$

1. Slice the distribution along lines $\theta \in S^{d-1}$
2. Project μ_1 and μ_2 onto θ: $P^\theta_\# \mu$, with $P^\theta : \mathbb{R}^d \to \mathbb{R}, \boldsymbol{x} \mapsto \langle \boldsymbol{x}, \theta \rangle$
3. Compute $\mathbb{R}^d$ Wasserstein onto the permutations obtained by sorting the projections
4. Average all the distances (mettre un theta en indice dans les sigma)

$$\mathcal{PWD}_2^2(\mu_1, \mu_2) \triangleq \int_{S^{d-1}} \frac{1}{n} \sum_{i=1}^{n} \left\| x_{\sigma_\theta(i)} - y_{\tau_\theta(i)} \right\|_2^2 d\omega(\theta),$$

with ω uniform distribution on S^{d-1}.



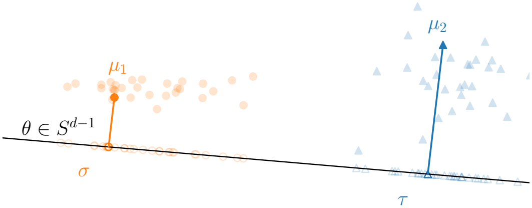
$\rightarrow$ provides an upper bound of $W_2^2(\mu_1, \mu_2)$ with complexity $O(Ln\, d + Ln \log(n))$, L number of lines

Sliced Wasserstein Generalized Geodesic
SWGG with a PWD-like formulation

1. Slice the distribution along lines $\theta \in S^{d-1}$
2. Project μ_1 and μ_2 onto θ: $P^\theta_\# \mu$, with $P^\theta : \mathbb{R}^d \to \mathbb{R}, \boldsymbol{x} \mapsto \langle \boldsymbol{x}, \theta \rangle$
3. Compute $\mathbb{R}^d$ Wasserstein onto the permutations obtained by sorting the projections
4. Take the minimum over all the distances

$$\mathrm{SWGG}_2^2(\mu_1, \mu_2, \theta) \triangleq \frac{1}{n} \sum_{i=1}^{n} \left\| x_{\sigma_\theta(i)} - y_{\tau_\theta(i)} \right\|_2^2,$$

$$\text{min-SWGG}_2^2(\mu_1, \mu_2) \triangleq \min_{\theta \in S^{d-1}} \mathrm{SWGG}_2^2(\mu_1, \mu_2, \theta)$$



Sliced Wasserstein Generalized Geodesic
SWGG with a PWD-like formulation

Properties of min-SWGG

- It comes with a **transport map**: let θ^* be the optimal projection direction

$$T(\boldsymbol{x}_i) = \boldsymbol{y}_{\tau_{\theta^*}^{-1}(\sigma_{\theta^*}(i))}, \quad \forall 1 \leq i \leq n.$$

- It is an upper bound of $\mathcal{W}$ and a lower bound of $\mathcal{PWD}$

$$\mathcal{W}_2^2 \leq \text{min-SWGG}_2^2 \leq \mathcal{PWD}_2^2$$

  and $\mathcal{W}_2^2 = \text{min-SWGG}_2^2$ when $d > 2n$ [2]
- Complexity $O(Lnd + Ln\log(n))$ with L number of lines
- The Monte-Carlo search over the L lines is effective in low dimension only

$\rightarrow$ **how to design gradient descent techniques for finding θ^*?**
$\rightarrow$ **further properties, such as sample complexity?**

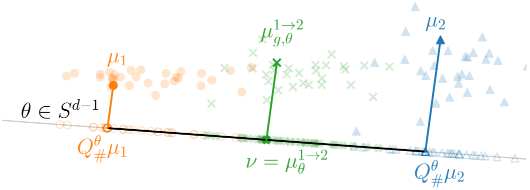
Sliced Wasserstein Generalized Geodesic
SWGG with a Generalized Geodesic formulation

1. Slice the distribution along lines $\theta \in S^{d-1}$
2. Project μ_1 and μ_2 onto θ: $Q^\theta_\# \mu$, with $Q^\theta : \mathbb{R}^d \to \mathbb{R}^d, \boldsymbol{x} \mapsto \theta \langle \boldsymbol{x}, \theta \rangle$
3. Define the pivot measure ν to be the Wasserstein mean of the measure $Q^\theta_\# \mu_1$ and $Q^\theta_\# \mu_2$

$$\nu = \mu_\theta^{1\to 2} \triangleq \arg\min_\mu \mathcal{W}_2^2(Q^\theta_\# \mu_1, \mu) + \mathcal{W}_2^2(\mu, Q^\theta_\# \mu_2)$$

4. Take the minimum over all the following distances

$$\text{SWGG}_2^2(\mu_1, \mu_2, \theta) = 2\mathcal{W}_2^2(\mu_1, \mu_\theta^{1\to 2}) + 2\mathcal{W}_2^2(\mu_\theta^{1\to 2}, \mu_2) - 4\mathcal{W}_2^2(\mu_{g,\theta}^{1\to 2}, \mu_\theta^{1\to 2})$$



$\to$ the two formulations are equivalent (for continuous or discrete distributions)

Sliced Wasserstein Generalized Geodesic
SWGG with a Generalized Geodesic formulation

Why this reformulation?

- Define a gradient descent algorithm for optimizing over θ
- Rewrite the problem as an OT formulation with a restricted constraint set
- Define new properties for SWGG

Properties of min-SWGG

- Weak convergence
- Translation invariance
- SWGG is equal to $\mathcal{W}$ when one of the distributions (μ_2) is supported on a line of direction θ:

$$\mathcal{W}_2^2(\mu_1, \mu_2) = \mathcal{W}_2^2(\mu_1, Q^\theta_\# \mu_1) + \mathcal{W}_2^2(Q^\theta_\# \mu_1, \mu_2)$$

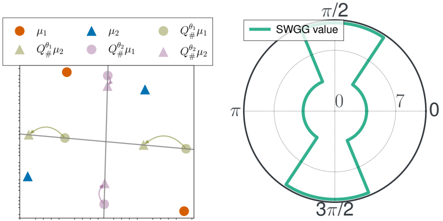
  that can be computed with a closed form

Sliced Wasserstein Generalized Geodesic

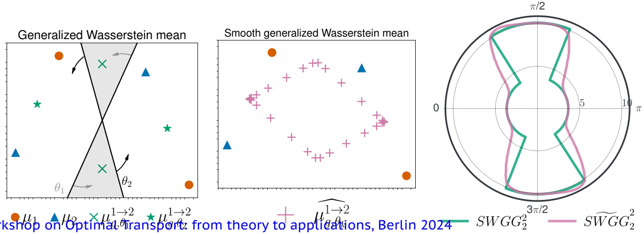
SWGG with a Generalized Geodesic formulation

Gradient descent for optimizing over θ

- min-$\text{SWGG}_2^2(\mu_1, \mu_2) = \min_{\theta \in \mathbb{S}^{d-1}} \frac{1}{n} \sum_{i=1}^{n} \left\| x_{\sigma_\theta(i)} - y_{\tau\theta(i)} \right\|_2^2$ is not amenable to optimization



- min-$\text{SWGG}_2^2(\mu_1, \mu_2) = \min_{\theta \in \mathbb{S}^{d-1}} 2\mathcal{W}_2^2(\mu_1, \mu_\theta^{1\to 2}) + 2\mathcal{W}_2^2(\mu_\theta^{1\to 2}, \mu_2) - 4\mathcal{W}_2^2(\mu_{g,\theta}^{1\to 2}, \mu_\theta^{1\to 2})$ can be computed with a $O(dn + n\log(n))$ complexity, but $\mathcal{W}_2^2(\mu_{g,\theta}^{1\to 2}, \mu_\theta^{1\to 2})$ is still piecewise linear with $\theta \to$ rely on the *blurred* Wasserstein distance [3]



Sliced Wasserstein Generalized Geodesic

SWGG with a Generalized Geodesic formulation

OT with a restricted constraint set

- Discrete optimal transport, with $n = m$ and uniform masses

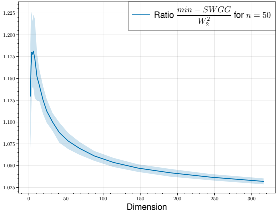
$$\mathcal{W}_2^2(\mu_1, \mu_2) \quad = \quad \min_{\gamma \in \Gamma(\mu_1, \mu_2)} \sum_{i,j} c(x_i, y_j)\gamma_{i,j}$$

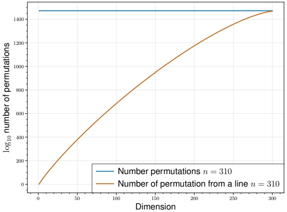
where $\Gamma(\mu_1, \mu_2) = \{\gamma \in \mathbb{R}^{n \times n} \text{ s.t. } \gamma \mathbb{1}_n = \mathbb{1}_n/n, \gamma^\top \mathbb{1}_n = \mathbb{1}_n/n\}$ (Birkhoff polytope).

- min-SWGG

$$\text{min-SWGG}_2^2(\mu_1, \mu_2) \quad = \quad \min_{\gamma_\theta \in \Pi(\mu_1, \mu_2)} \sum_{i,j} c(x_i, y_j)\gamma_{\theta\, i,j}$$

where $\Pi(\mu_1, \mu_2) = \{\gamma_\theta \in \mathbb{R}^{n \times n}$ s.t. it is constructed from the permutahedron of the proj. distributions$\}$





Sliced Wasserstein Generalized Geodesic
SWGG with a Generalized Geodesic formulation

OT with a restricted constraint set

- Discrete optimal transport, with $n = m$ and uniform masses

$$\mathcal{W}_2^2(\mu_1, \mu_2) \quad = \quad \min_{\gamma \in \Gamma(\mu_1,\mu_2)} \sum_{i,j} c(x_i, y_j)\gamma_{i,j}$$

  where $\Gamma(\mu_1, \mu_2) = \{\gamma \in \mathbb{R}^{n\times n} \text{ s.t. } \gamma 1_n = 1_n/n, \gamma^\top 1_n = 1_n/n\}$ (Birkhoff polytope).

- min-SWGG

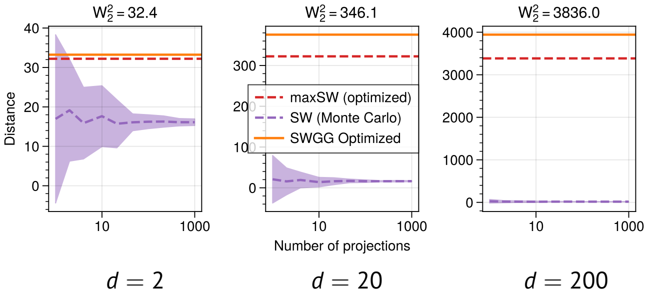
$$\text{min-SWGG}_2^2(\mu_1, \mu_2) \quad = \quad \min_{\gamma_\theta \in \Pi(\mu_1,\mu_2)} \sum_{i,j} c(x_i, y_j)\gamma_{\theta\, i,j}$$

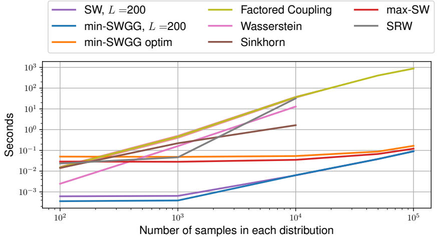
  where $\Pi(\mu_1, \mu_2) = \{\gamma_\theta \in \mathbb{R}^{n\times n}$ s.t. it is constructed from the permutahedron of the proj. distributions$\}$

- $\Pi(\mu_1, \mu_2) \subset \Gamma(\mu_1, \mu_2)$
- Gives a sample complexity similar to Sinkhorn $n^{-1/2}$ *measures lying on smaller dimensional subspaces has a better sample complexity than between the original measures*

Experimental results
Computational aspects

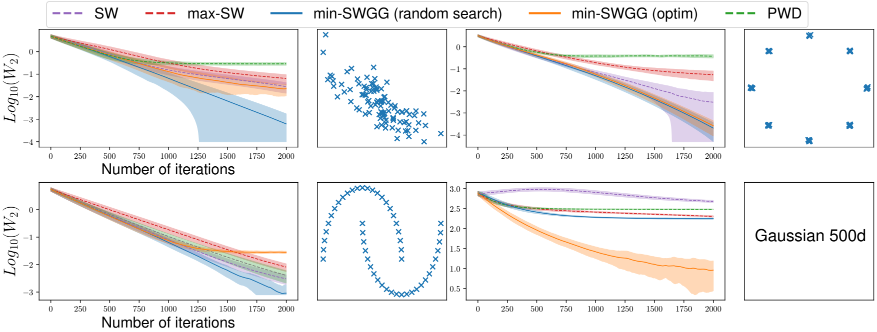
- Two Gaussian distributions μ_1 and μ_2





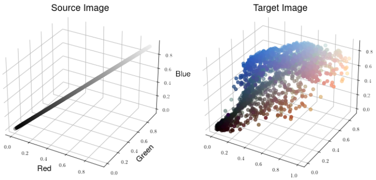
Experimental results
Gradient flows

- Initial μ_1: uniform distribution, different target distributions



Experimental results
Pan sharpening / image colorization, using the map

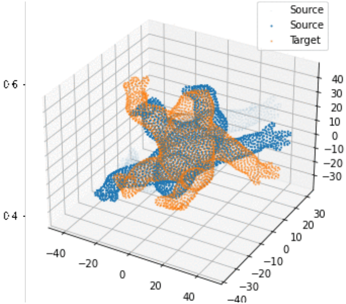
- One distribution is supported on a line



- Construct a super-resolution multi-chromatic satellite image from a high-resolution mono-chromatic image (source) and low-resolution multi-chromatic image (target)

Experimental results
Point cloud matchings, using the map

- Iterative Closest Point iterative algorithm for aligning point clouds
- Based on several one-to-one correspondences between points



n	500	3000	150 000
NN	3.54 (**0.02**)	96.9 (**0.30**)	23.3 (**59.37**)
OT	0.32 (0.18)	48.4 (58.46)	·
min-SWGG	**0.05** (0.04)	**37.6** (0.90)	**6.7** (105.75)

(the lower the better, timings into parenthesis)

Experimental results
Optimal transport dataset distances

- For computing distances between datasets
- Cumbersome to compute in practice since it lays down on solving multiple OT problems

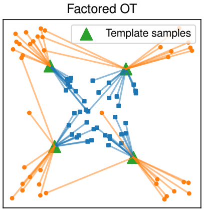
	MNIST	EMNIST	Fashion	KMNIST	USPS
MNIST		1	1.3	1.2	1
EMNIST	1		1.2	1.2	1
Fashion	1.3	1.2		1.3	0.8
KMNIST	1.2	1.2	1.3		1
USPS	1	1	0.8	1.1	

	MNIST	EMNIST	Fashion	KMNIST	USPS
MNIST		0.9	1.2	1.1	0.9
EMNIST	0.8		1.1	1.1	0.9
Fashion	1.2	1.3		1.2	0.8
KMNIST	1.1	1	1.2		0.9
USPS	0.9	0.9	0.7	0.9	

Figure: OTDD results ($\times 10^2$) distances for min-SWGG (left) and Sinkhorn divergence (right) for various datasets.

Conclusion

- Sliced Wasserstein Generalized Geodesic
  - provides an upper bound for Wasserstein
  - comes with an associated transport map
  - has a $O(Lnd + n\log(n))$ complexity
  - has good statistical properties
- Not the only approximation method based on a pivot measure
  - Factored coupling [4], where $\nu = \arg\min_{\mu \in \mathcal{P}(\mathbb{R}^k)} \{\mathcal{W}_2^2(\mu, \mu_1) + \mathcal{W}_2^2(\mu, \mu_1)\}$



  - Subspace detours [6], where $\nu = \arg\min_{\nu \in \mathcal{P}(\mathbb{R}^d)} \left\{\mathcal{W}_2^2(P^E_{\#}\mu_1, \nu) + \mathcal{W}_2^2(\nu, P^E_{\#}\mu_2)\right\}$
- Some open questions
  - how do the Birkhoff polytope and the considered permutahedron relate?
  - concentration results?
  - extension to incomparable spaces through a pivot measure?

Fast Optimal Transport through Sliced Generalized Wasserstein Geodesics

Joint work with Guillaume Mahey, Gilles Gasso, Clément Bonet and Nicolas Courty
NeurIPS 2023 [5]

Laetitia Chapel
laetitia.chapel@irisa.fr

IRISA, Rennes, France

Institut Agro Rennes-Angers

Workshop on Optimal Transport: from theory to applications, Berlin 2024

Bibliography I

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2005.

[2] Thomas M Cover. "The number of linearly inducible orderings of points in d-space". In: *SIAM Journal on Applied Mathematics* 15.2 (1967), pp. 434–439.

[3] Jean Feydy. "Geometric data analysis, beyond convolutions". PhD thesis. École Normale Supérieure de Cachan, 2020.

[4] Aden Forrow et al. "Statistical optimal transport via factored couplings". In: *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR. 2019, pp. 2454–2465.

[5] Guillaume Mahey et al. "Fast Optimal Transport through Sliced Generalized Wasserstein Geodesics". In: *Advances in Neural Information Processing Systems* 36 (2024).

[6] Boris Muzellec and Marco Cuturi. "Subspace detours: Building transport plans that are optimal on subspace projections". In: *Advances in Neural Information Processing Systems* 32 (2019).