Sliced optimal partial transport

Yikun Bai[1a] Bernhard Schmitzer[1b] Mathew Thorpe[c]
Soheil Kolouri[a]

[a]Department of Computer Science, Vanderbilt University

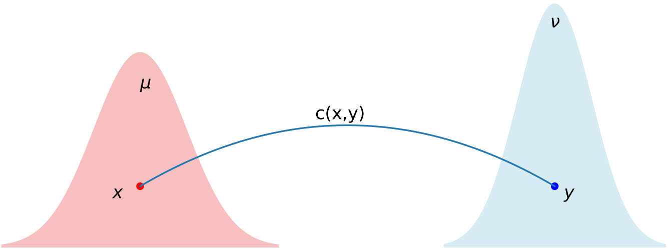
[b]Institude of Computer Science, Gottingen University

[c]The Allen Turing Institute

May 27, 2023

[1]These authors have equal contribution.

Introduction: optimal transport problem (OT)

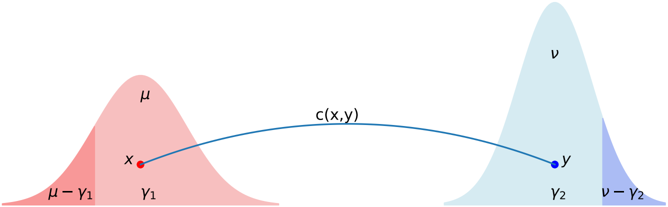


$$\mathrm{OT}(\mu,\nu) := \inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\Omega^2} c(x,y)d\gamma(x,y)$$

where $\Omega \subset \mathbb{R}^d$, $\mu,\nu \in \mathcal{P}(\Omega), \Gamma(\mu,\nu) := \{\gamma \in \mathcal{P}(\Omega^2) : \gamma_1 = \mu, \gamma_2 = \nu\}$

- Statistics: hypothesis test, statistical inference
- Machine learning: GAN, VAE, transfer learning
- Limitation: Requires equal total amount of mass between the two measures.

Introduction: optimal partial transport (OPT)



$$\text{OPT}_\lambda(\mu,\nu) := \inf_{\gamma\in\Gamma_{\leq}(\mu,\nu)} \int_{\Omega^2} c(x,y)d\gamma(x,y) + \lambda(|\mu-\gamma_1|_{TV} + |\nu-\gamma_2|_{TV}).$$

where $\mu,\nu \in \mathcal{M}_+(\Omega)$, $\Gamma_{\leq}(\mu,\nu) := \{\gamma \in \mathcal{M}_+(\Omega^2) : \gamma_1 \leq \mu, \gamma_2 \leq \nu\}$, and $\lambda \geq 0$.

- **Benefits**: Partial matching and comparison of measures with unequal mass.

Introduction: Empirical OPT

If $\mu = \sum_{i=1}^{n} \delta_{x_i}, \nu = \sum_{j=1}^{m} \delta_{y_j}$, OPT problem becomes

$$OPT_\lambda(\mu, \nu) = \inf_{\gamma \in \Gamma_\leq(\mu,\nu)} \sum_{i,j} c(x_i, y_j)\gamma_{ij} + \lambda(m + n - 2|\gamma|)$$

where $\Gamma_\leq(1_n, 1_m) := \{\gamma \in \mathbb{R}_+^{n\times m} : \gamma 1_m \leq 1_n, \gamma^T 1_n \leq 1_m\}, |\gamma| = \sum_{ij} \gamma_{ij}$.
Challenge:

- High dimension linear programming problem

Existing methods:

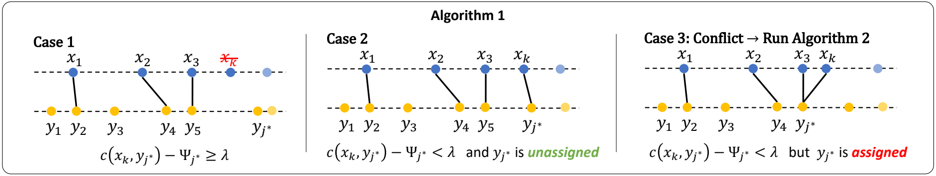
- Network simplex (Bonneel et al. 2011): $\mathcal{O}((n + m)nm)$
- Sinkhorn algorithm (Chizat et al. 2018): $\mathcal{O}(\frac{1}{\epsilon^3} nm)$, where ϵ is weight of entropic regularization
- Dynamic programming (Sato et al. 2020): $\mathcal{O}(\ln^2(n + m)(n + m))$. Requires tree metric assumption.

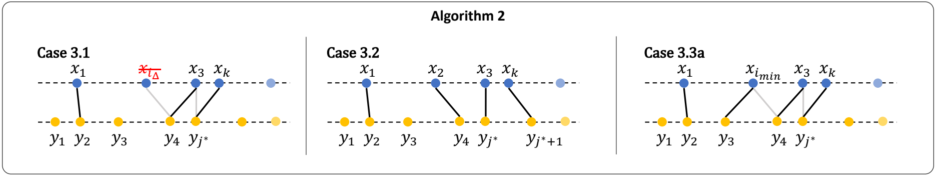
Our work: Solve the 1D OPT

In $\mathbb{R}$, consider the empirical OPT problem can be simplified to the following linear alignment problem [Bai et al. 2022, Proposition 3.1]

$$\text{OPT}_\lambda(\mu,\nu) = \min_L \sum_{i \in \text{Dom}(L)} c(x_i, y_{L[j]}) + \lambda(n + m - 2|\text{dom}(L)|)$$

where $L : [1 : n] \to [1 : m]$ is partial increasing bijection.

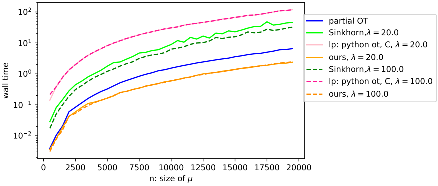




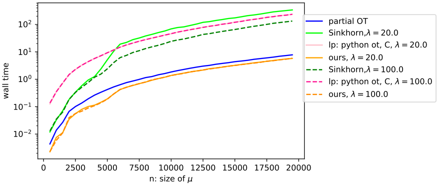
Accuracy and time complexity of the main algorithm

By [Bai et al. 2022, Theorem 4.4]:

- (Accuracy) Algorithm 1 solves the empirical OPT problem.
- (Time complexity) In the worst case, the time complexity of algorithm 1 is $\mathcal{O}(n \max(n, m))$.



(a) uniform distributions



(b) Gaussian mixture distributions

Sliced optimal partial transport (SOPT)

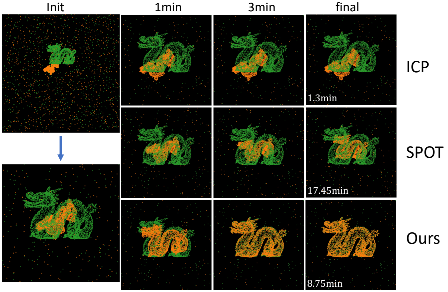
$$SOPT_\lambda(\mu,\nu) := \int_{\mathbb{S}^{d-1}} \mathrm{OPT}_{\lambda(\theta)}(\langle\theta,\cdot\rangle_{\#}\mu, \langle\theta,\cdot\rangle_{\#}\nu)d\sigma(\theta)$$
$$\approx \frac{1}{T}\sum_{t=1}^{T} \mathrm{OPT}_{\lambda_t}(\langle\theta,\cdot\rangle_{\#}\mu, \langle\theta,\cdot\rangle_{\#}\nu)d\sigma(\theta)$$

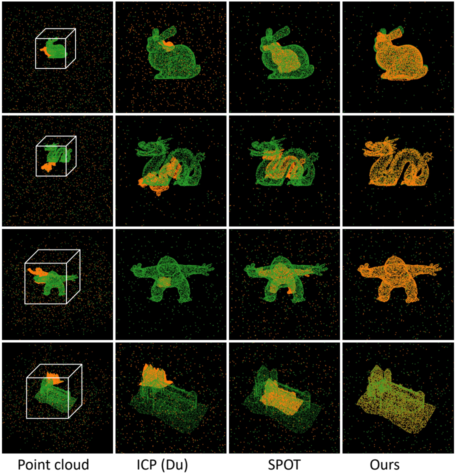
where $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\|^2 = 1\}$, $\sigma = \mathrm{Unif}(\mathbb{S}^{d-1})$, $\mathrm{supp}(\sigma) = \mathbb{S}^{d-1}$, $\lambda : \mathbb{S}^{d-1} \to \mathbb{R}_{++}$ is L_1 function, and $f_{\#}\mu$ is the push-forward measure of μ for any (measurable) f.

- $SOPT$ is a metric in $\mathcal{M}_+(\Omega)$ if c is a metric.
- $SOPT$ can be regarded as a proxy of $OPT_\lambda(\mu,\nu)$ distance.

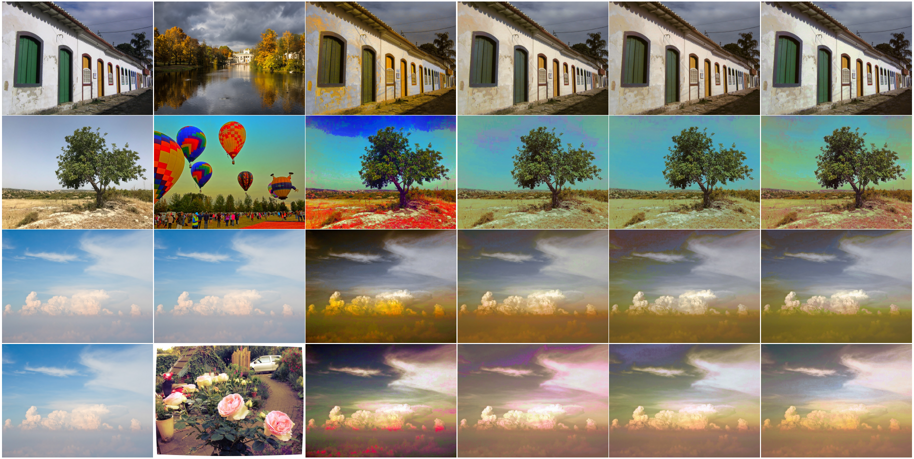
Experiment: Noisy Point Cloud Registration

In $\mathbb{R}^3$, given two measure (μ, ν) with $\nu = T_{\#}\mu$ where mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ has form: $T(x) = sRx + \beta$, R is rotation matrix, scaling $s > 0$, translation $\beta \in \mathbb{R}^3$. Given samples with noise corruption, how to estimate T?





Experiment: Color adaptation



Source	target	OT	SPOT	ours($\lambda = 10$)	ours($\lambda \leq 2$)

Summary

Our contributions:

- We proposed a new quadratic time algorithm for 1D OPT problem.
- We propose the so called sliced-optimal partial transport distance (SOPT).
- We demonstrate the applications of SOPT in point cloud registration and color adaptation.

Future's work:

- Potential applications of SOPT in GAN, VAE

Reference

- Bai, Yikun et al. (2022). "Sliced optimal partial transport". In: *arXiv preprint arXiv:2212.08049*.
- Bonneel, Nicolas et al. (2011). "Displacement Interpolation Using Lagrangian Mass Transport". In: *ACM Transactions on Graphics (SIGGRAPH ASIA 2011)* 30.6.
- Chizat, Lenaic et al. (2018). "Scaling algorithms for unbalanced optimal transport problems". In: *Mathematics of Computation* 87.314, pp. 2563–2609.
- Sato, Ryoma et al. (2020). "Fast and robust comparison of probability measures in heterogeneous spaces". In: *arXiv preprint arXiv:2002.01615*.

Thank you