# Subspace Detours Meet Gromov-Wasserstein

Clément Bonet[1], Nicolas Courty[1], François Septier[1], Lucas Drumetz[2]

[1]Université Bretagne Sud
[2]IMT Atlantique

OTML 2021
13/12/2021

# Gromov-Wasserstein Distance

Let $(X, d_X, \mu)$, $(Y, d_Y, \nu)$ be metric-measures spaces (mm-spaces).

## Gromov-Wasserstein

$$GW(X,Y) = \inf_{\gamma \in \Pi(\mu,\nu)} \iint \big(c_X(x,x') - c_Y(y,y')\big)^2 \,\mathrm{d}\gamma(x,y)\mathrm{d}\gamma(x',y')$$

with $\Pi(\mu,\nu) = \{\gamma \in \mathcal{P}(X \times Y),\ \pi^1_\# \gamma = \mu,\ \pi^2_\# \gamma = \nu\}$.

## Examples

For $X = \mathbb{R}^p$, $Y = \mathbb{R}^q$,

- $c_X(x,x') = d_X(x,x')$, $c_Y(y,y') = d_Y(y,y)$
- $c_X(x,x') = \|x - x'\|_2^2$, $c_Y(y,y') = \|y - y'\|_2^2$
- $c_X(x,x') = \langle x, x' \rangle_p$, $c_Y(y,y') = \langle y, y' \rangle_q$

Properties:

- Distance between mm-spaces up to isometries
- Invariances

# Gromov-Wasserstein Distance

Let $(X, d_X, \mu)$, $(Y, d_Y, \nu)$ be metric-measures spaces (mm-spaces).

## Gromov-Wasserstein

$$GW(X,Y) = \inf_{\gamma \in \Pi(\mu,\nu)} \iint \left(c_X(x,x') - c_Y(y,y')\right)^2 \, \mathrm{d}\gamma(x,y)\mathrm{d}\gamma(x',y')$$

with $\Pi(\mu,\nu) = \{\gamma \in \mathcal{P}(X \times Y),\ \pi^1_\# \gamma = \mu,\ \pi^2_\# \gamma = \nu\}$.

Practical Issues:

- Computational Complexity: $O(n^4)$

Proposed Solution:

- Entropic Regularization [Peyré et al., 2016]
- Low rank constraints [Scetbon et al., 2021]
- Sliced Gromov-Wasserstein [Vayer et al., 2019]
- Minibatch estimators [Fatras et al., 2021]

# Subspace Detour Approach

Subspace Detour approach [Muzellec and Cuturi, 2019]:

- Choose a subspace $E$
- Project the measures: $\mu_E = \pi^E_\# \mu$ and $\nu_E = \pi^E_\# \nu$
- Take the optimal coupling $\gamma^*_E \in \Pi(\mu_E, \nu_E)$
- Find a coupling $\gamma \in \Pi_E(\mu, \nu) = \{\gamma \in \Pi(\mu, \nu) |\ (\pi^E, \pi^E)_\# \gamma = \gamma^*_E\}$

Coupling on the whole set:

- Monge-Independent plan:

$$\pi_{\mathrm{MI}} = \gamma^*_E \otimes (\mu_{E^\perp|E} \otimes \nu_{E^\perp|E})$$

- Monge-Knothe plan:

$$\pi_{\mathrm{MK}} = \gamma^*_E \otimes \gamma^*_{E^\perp|E}$$

# Motivation

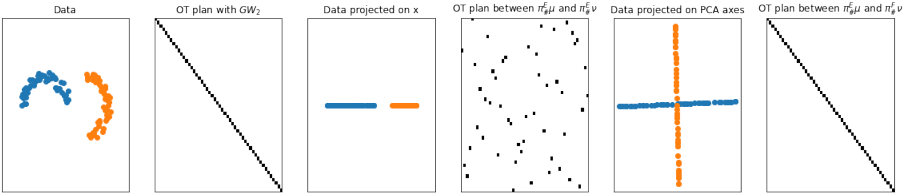



Figure: From left to right: Data (moons), OT plan obtained with GW for $c(x, x') = \|x - x'\|_2^2$, Data projected on the 1st axis, OT plan obtained between the projected measures, Data projected on their 1st PCA component, OT plan obtained between the the projected measures.

Subspace Detour approach for GW: Let $\mu \in \mathcal{P}(\mathbb{R}^p)$, $\nu \in \mathcal{P}(\mathbb{R}^q)$,

- Choose subspace $E \subset \mathbb{R}^p$ and $F \subset \mathbb{R}^q$
- Project the measures: $\mu_E = \pi^E_\# \mu$ and $\nu_F = \pi^F_\# \nu$
- Take the optimal coupling $\gamma^*_{E,F} \in \Pi(\mu_E, \nu_F)$
- Find a coupling $\gamma \in \Pi_{E,F}(\mu, \nu) = \{\gamma \in \Pi(\mu, \nu) | \ (\pi^E, \pi^F)_\# \gamma = \gamma^*_{E \times F}\}$

# Properties

$$GW_{E,F}(\mu,\nu) = \inf_{\gamma \in \Pi_{E,F}(\mu,\nu)} \iint \left(c_X(x,x') - c_Y(y,y')\right)^2 \,\mathrm{d}\gamma(x,y)\mathrm{d}\gamma(x',y')$$

Properties:

- Let $\pi_{\mathrm{MK}} = \gamma^*_{E\times F} \otimes \gamma^*_{E^\perp \times F^\perp | E\times F}$, then

$$\pi_{\mathrm{MK}} \in \operatorname*{argmin}_{\gamma \in \Pi_{E,F}(\mu,\nu)} \iint \left(c_X(x,x') - c_Y(y,y')\right)^2 \,\mathrm{d}\gamma(x,y)\mathrm{d}\gamma(x',y').$$

- $GW_{E,F}$ invariant *w.r.t* isometries of the form $f = (Id_E, f_{E^\perp})$ for $c_X(x,x') = \|x - x'\|_2^2$, $c_Y(y,y') = \|y - y'\|_2^2$ or $c_X(x,x') = \langle x, x'\rangle_p$, $c_Y(y,y') = \langle y, y'\rangle_q$.

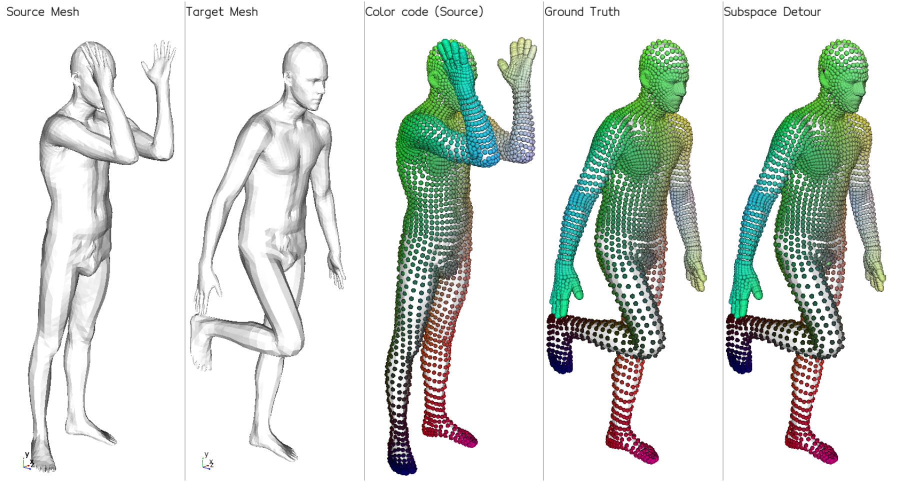
# Choice of subspace

- Better sample complexity on lower dimensional spaces [Vayer et al., 2018]
- Better computational cost with 1D subspaces: $O(n \log n)$ [Vayer et al., 2019]

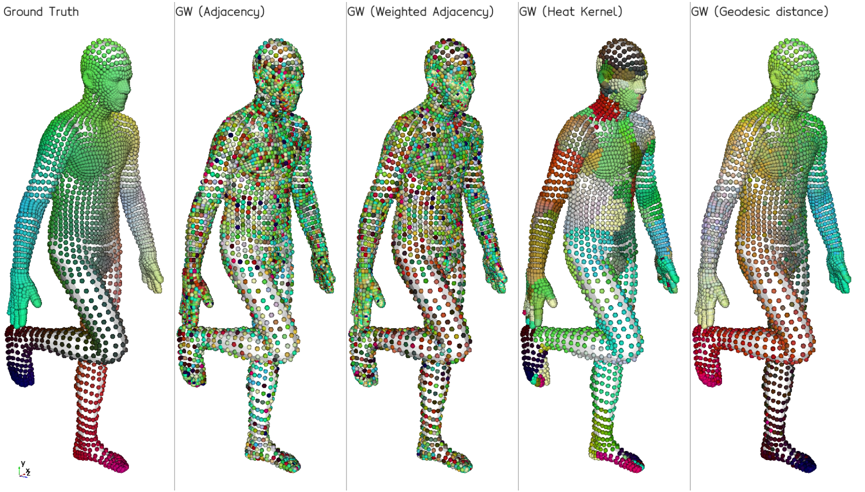
Choice of subspace?

- Euclidean spaces: PCA
- Graphs: Fiedler vector
- Gradient descent over Stiefel manifold

# Application



# Application



# Conclusion

Thank you!

Paper: https://arxiv.org/abs/2110.10932

# References I

Kilian Fatras, Younes Zine, Szymon Majewski, Rémi Flamary, Rémi Gribonval, and Nicolas Courty. Minibatch optimal transport distances; analysis and applications. *arXiv preprint arXiv:2101.01792*, 2021.

Boris Muzellec and Marco Cuturi. Subspace detours: Building transport plans that are optimal on subspace projections. *arXiv preprint arXiv:1905.10099*, 2019.

Gabriel Peyré, Marco Cuturi, and Justin Solomon. Gromov-wasserstein averaging of kernel and distance matrices. In *International Conference on Machine Learning*, pages 2664–2672. PMLR, 2016.

Meyer Scetbon, Marco Cuturi, and Gabriel Peyré. Low-rank sinkhorn factorization. *arXiv preprint arXiv:2103.04737*, 2021.

Titouan Vayer, Laetita Chapel, Rémi Flamary, Romain Tavenard, and Nicolas Courty. Fused gromov-wasserstein distance for structured objects: theoretical foundations and mathematical properties. *arXiv preprint arXiv:1811.02834*, 2018.

Titouan Vayer, Rémi Flamary, Nicolas Courty, Romain Tavenard, and Laetitia Chapel. Sliced gromov-wasserstein. *Advances in Neural Information Processing Systems*, 32:14753–14763, 2019.