On the existence of Monge maps for the Gromov–Wasserstein problem

Joint work with T. Dumont and F.-X. Vialard.

Journée Transport Optimal et ses applications (GdR IASIS)
17 février 2025

theo.lacombe@univ-eiffel.fr

The Gromov–Wasserstein Problem

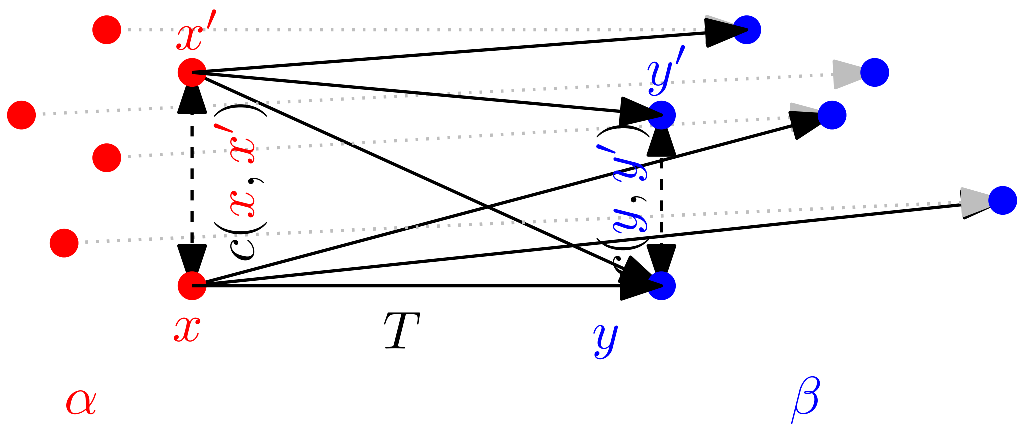
Definition:

Let $X, Y \subset \mathbb{R}^d$, $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ be a cost function, and α, β be two probability measures supported on X, Y respectively. The Gromov–Wasserstein problem is defined as

$$\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x, x') - c(y, y')|^2 \mathrm{d}\pi(x, y)\mathrm{d}\pi(x', y'), \tag{GW}$$

where $\Pi(\alpha, \beta)$ denote the set of transport plans between α and β, i.e. measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals α and β.

Example: $\alpha = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\beta = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$ ("point clouds"). Then $\pi \simeq$ a bistochastic matrix of size $n \times n$ ("bipartite graph matching"), i.e. $\pi_{ij} \leftrightarrow$ the mass transported from x_i to y_j



In a nutshell:

GW measures an "isometry" defect (for the cost c) between X, α and Y, β.

The Gromov–Wasserstein Problem

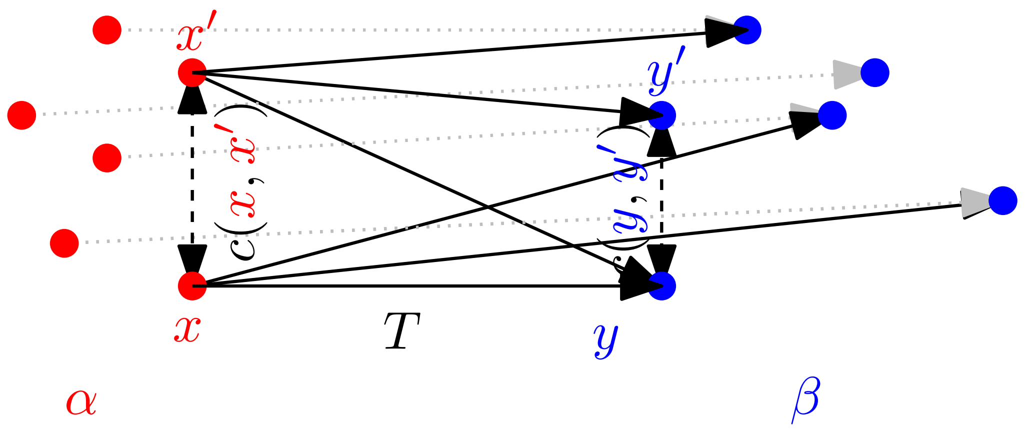
Definition:

Let $X, Y \subset \mathbb{R}^d$, $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ be a cost function, and α, β be two probability measures supported on X, Y respectively. The Gromov–Wasserstein problem is defined as

$$\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x, x') - c(y, y')|^2 \mathrm{d}\pi(x, y) \mathrm{d}\pi(x', y'), \qquad \text{(GW)}$$

where $\Pi(\alpha, \beta)$ denote the set of transport plans between α and β, i.e. measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals α and β.

Example: $\alpha = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\beta = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$ ("point clouds"). Then $\pi \simeq$ a bistochastic matrix of size $n \times n$ ("bipartite graph matching"), i.e. $\pi_{ij} \leftrightarrow$ the mass transported from x_i to y_j



Question 1: Under which conditions are the solution of (GW) *deterministic*, that is of the form $(\mathrm{id}, T)_\# \alpha$ for some measurable map T?

The Gromov–Wasserstein Problem

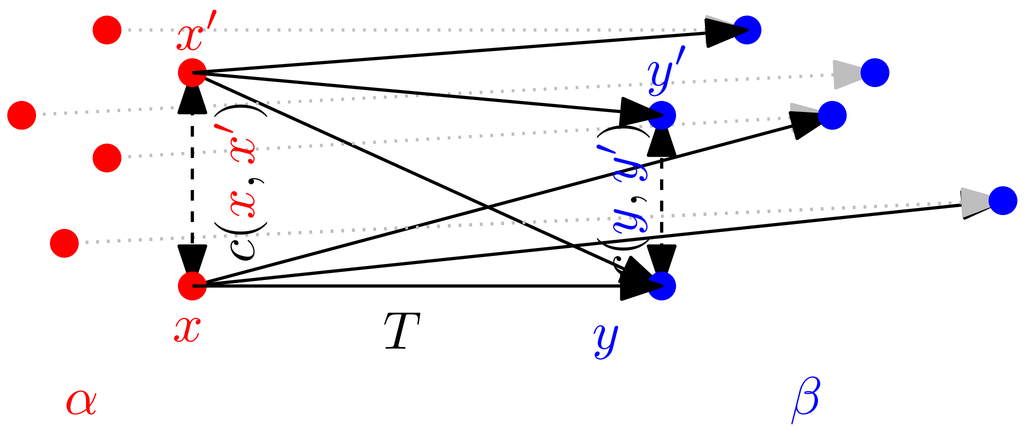
Definition:

Let $X, Y \subset \mathbb{R}^d$, $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ be a cost function, and α, β be two probability measures supported on X, Y respectively. The Gromov–Wasserstein problem is defined as

$$\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y'), \tag{GW}$$

where $\Pi(\alpha, \beta)$ denote the set of transport plans between α and β, i.e. measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals α and β.

Example: $\alpha = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\beta = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$ ("point clouds"). Then $\pi \simeq$ a bistochastic matrix of size $n \times n$ ("bipartite graph matching"), i.e. $\pi_{ij} \leftrightarrow$ the mass transported from x_i to y_j



Question 1: Under which conditions are the solution of (GW) *deterministic*, that is of the form $(\mathrm{id}, T)_\# \alpha$ for some measurable map T?

Question 2: Are there instances where GW is easy to compute? e.g. if $d = 1$?

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot,\cdot\rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha,\beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha,\beta)$ **if and only if** it minimizes

$$\Pi(\alpha,\beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y')\right)}_{=:C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi\rangle\,. \tag{OT}$$

Idea of proof: In general, introduce the bilinear relaxation $F : (\pi,\gamma) \mapsto \iint k \mathrm{d}\pi\otimes\gamma = \langle \pi, k\gamma\rangle$, with k symmetric, bilinear and negative on $\mathrm{span}(\{\pi - \gamma, \pi,\gamma\in\Pi(\alpha,\beta)\})$.
Then, let $(\pi^\star,\gamma^\star)$ minimize F (in particular, $\gamma^\star$ minimizes $\pi\mapsto F(\pi^\star,\pi)$), and observe that it yields $\langle \pi^\star - \gamma^\star, k(\pi^\star-\gamma^\star)\rangle \geqslant 0$, hence $= 0$, hence $(\pi^\star - \gamma^\star) \in \ker(k)$, and thus $F(\pi^\star,\pi^\star) = F(\pi^\star,\gamma^\star)$.
It means that if $\pi^\star$ minimizes $\pi\mapsto F(\pi,\pi)$, it also minimizes $F(\pi^\star,\pi)$, and vice-versa.
It turns out that $k(x,x',y,y') = |c(x,x') - c(y,y')|^2$ is indeed negative (on measures of mass 0) for the two costs considered in this work.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot,\cdot\rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha,\beta)$. Then, any $\tilde\pi^\star$ minimizes $\mathrm{GW}(\alpha,\beta)$ **if and only if** it minimizes

$$\Pi(\alpha,\beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y')\right)}_{=:C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi\rangle . \quad \text{(OT)}$$

Observation: This problem is linear in π. It is the Optimal Transportation problem between α and β for the (unknown...) cost $C_{\pi^\star}$. Thanks to [Brenier 1987, Villani 2008, McCann 2011, Moameni 2016], we know that if α has a Lebesgue density, we know sufficient conditions on the cost $C_{\pi^\star}$ to ensure that minimizers of (OT) are induced by transport maps. Namely:

Proposition: [Moameni 2016]

If α has a density and $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies

$$\forall x_0, y_0,\ |\{y |\ \nabla_x C(x_0,y) = \nabla_x C(x_0,y_0)\}| \leqslant m, \quad (m\text{-twist})$$

then any minimizer of $\pi \mapsto \langle C,\pi\rangle$ in $\Pi(\alpha,\beta)$ is supported on the union of (at most) m transport maps.

In a nutshell:

If $\nabla_x C$ is "m-injective", each x is sent to (at most) m y.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot, \cdot \rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha,\beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha,\beta)$ **if and only if** it minimizes

$$\Pi(\alpha,\beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y') \right)}_{=: C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi \rangle\,. \quad \text{(OT)}$$

Observation: This problem is linear in π. It is the Optimal Transportation problem between α and β for the (unknown...) cost $C_{\pi^\star}$. Thanks to [Brenier 1987, Villani 2008, McCann 2011, Moameni 2016], we know that if α has a Lebesgue density, we know sufficient conditions on the cost $C_{\pi^\star}$ to ensure that minimizers of (OT) are induced by transport maps. Namely:

Proposition: [Moameni 2016]

If α has a density and $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies

$$\forall x_0, y_0, \ |\{y | \ \nabla_x C(x_0,y) = \nabla_x C(x_0,y_0)\}| \leqslant m, \quad (m\text{-twist})$$

then any minimizer of $\pi \mapsto \langle C, \pi \rangle$ in $\Pi(\alpha,\beta)$ is supported on the union of (at most) m transport maps.

In a nutshell:

If $\nabla_x C$ is "m-injective", each x is sent to (at most) m y.

Question: does $C_{\pi^\star}$ satisfy the m-twist condition?

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x, x') - c(y, y')|^2 \mathrm{d}\pi(x, y)\mathrm{d}\pi(x', y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot, \cdot \rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha, \beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha, \beta)$ **if and only if** it minimizes

$$\Pi(\alpha, \beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x, x') - c(y, y')|^2 \mathrm{d}\pi^\star(x', y') \right)}_{=: C_{\pi^\star}(x,y)} \mathrm{d}\pi(x, y) = \langle C_{\pi^\star}, \pi \rangle . \quad \text{(OT)}$$

Observation: This problem is linear in π. It is the Optimal Transportation problem between α and β for the (unknown...) cost $C_{\pi^\star}$. Thanks to [Brenier 1987, Villani 2008, McCann 2011, Moameni 2016], we know that if α has a Lebesgue density, we know sufficient conditions on the cost $C_{\pi^\star}$ to ensure that minimizers of (OT) are induced by transport maps. Namely:

Proposition: [Moameni 2016]

If α has a density and $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies

$$\forall x_0, y_0, \ |\{y | \ \nabla_x C(x_0, y) = \nabla_x C(x_0, y_0)\}| \leqslant m, \quad (m\text{-twist})$$

then any minimizer of $\pi \mapsto \langle C, \pi \rangle$ in $\Pi(\alpha, \beta)$ is supported on the union of (at most) m transport maps.

In a nutshell:

If $\nabla_x C$ is "m-injective", each x is sent to (at most) m y.

Question: does $C_{\pi^\star}$ satisfy the m-twist condition?

Answer: No (in general).

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot,\cdot\rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha,\beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha,\beta)$ **if and only if** it minimizes

$$\Pi(\alpha,\beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y')\right)}_{=:C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi\rangle \,. \qquad \text{(OT)}$$

Inner-product case: if $c = \langle\cdot,\cdot\rangle$, observe that

$$\text{(GW)} \Leftrightarrow \min_\pi - \iint \langle x,x'\rangle \langle y,y'\rangle \mathrm{d}\pi^\star(x',y')\mathrm{d}\pi(x,y) \Rightarrow \text{ wlog } \; C_{\pi^\star}(x,y) = -\langle M^\star x, y\rangle$$

with $M^\star := \int x' y'^T \mathrm{d}\pi^\star(x',y')$.

If $M^\star$ is of full rank, $\nabla_x C_{\pi^\star}(x,y) = -(M^\star)^T y$ that is injective $\Rightarrow$ 1-twist condition $\Rightarrow \exists$ optimal transport maps![a] Otherwise...?

Idea: Up to a Singular Value Decomposition, $M^\star$ is a projection and, *on the image of the projection*, the 1-twist condition is satisfied so we get the existence of a transport map. Can we lift it?

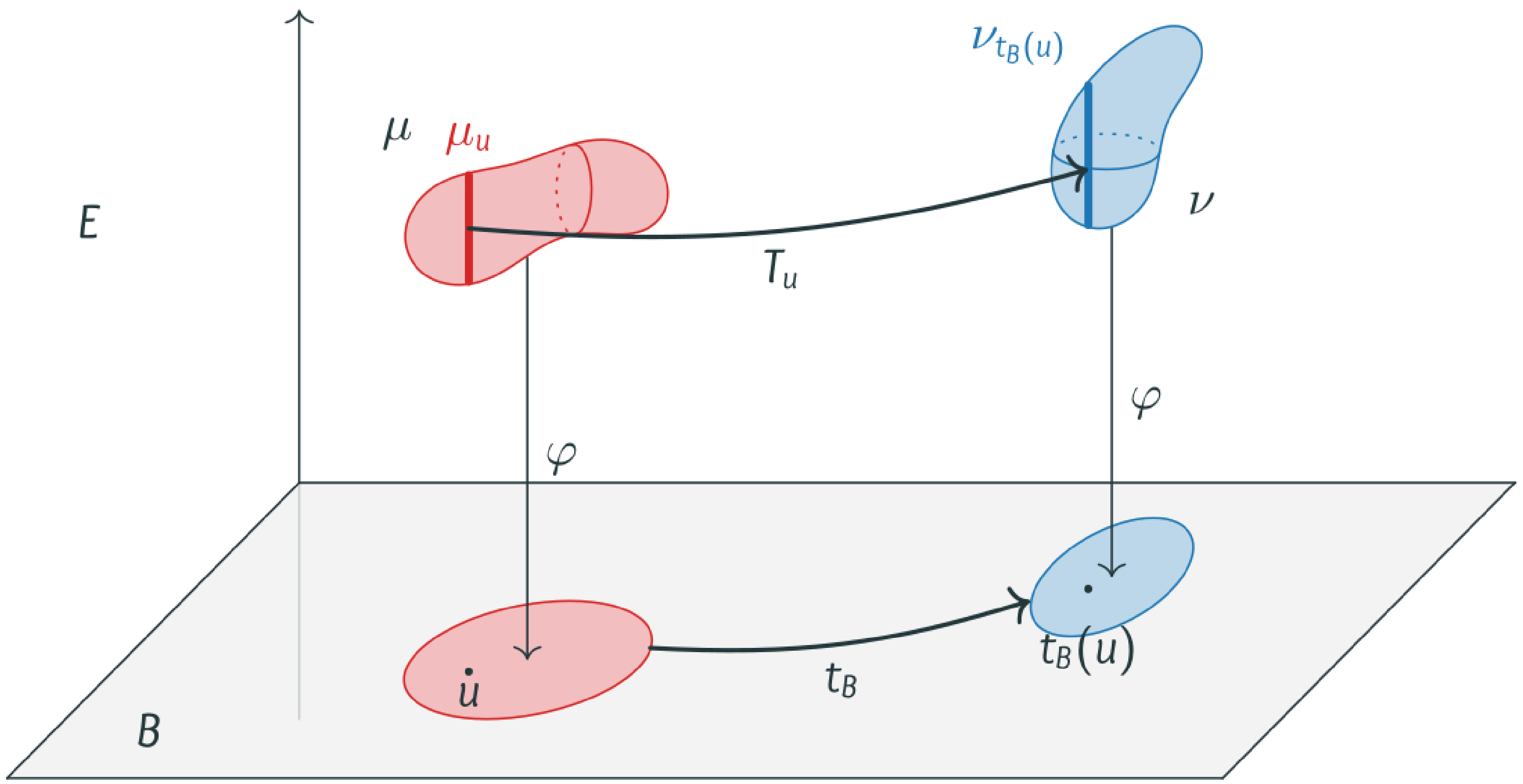
[a] Already proved by Vayer et al.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x, x') - c(y, y')|^2 \mathrm{d}\pi(x, y) \mathrm{d}\pi(x', y')$.

Proposition: [Dumont, L, Vialard, 2024], "Brenier theorem for fiber-invariant costs", Informal

Let $\alpha, \beta \in \mathcal{P}(E)$, κ a cost function, and $\varphi : E \to B$ such that (i) $\varphi\#\mu \ll \mathrm{Vol}(B)$, (ii) $\varphi\#\alpha$-a.e., the fiber $\varphi^{-1}(u)$ is a complete manifold (iii) $\varphi\#\alpha$-a.e., the disintegration α_u of α wrt φ is $\ll \mathrm{Vol}(\varphi^{-1}(u))$, (iv) $\kappa(x, y) = \tilde{\kappa}(\varphi(x), \varphi(y))$ with $\tilde{\kappa}$ twisted on B. Then $\exists$ an optimal transport map between α and β (with some "gradient of convex function" structure on t_B and all $(T_u)_u$).



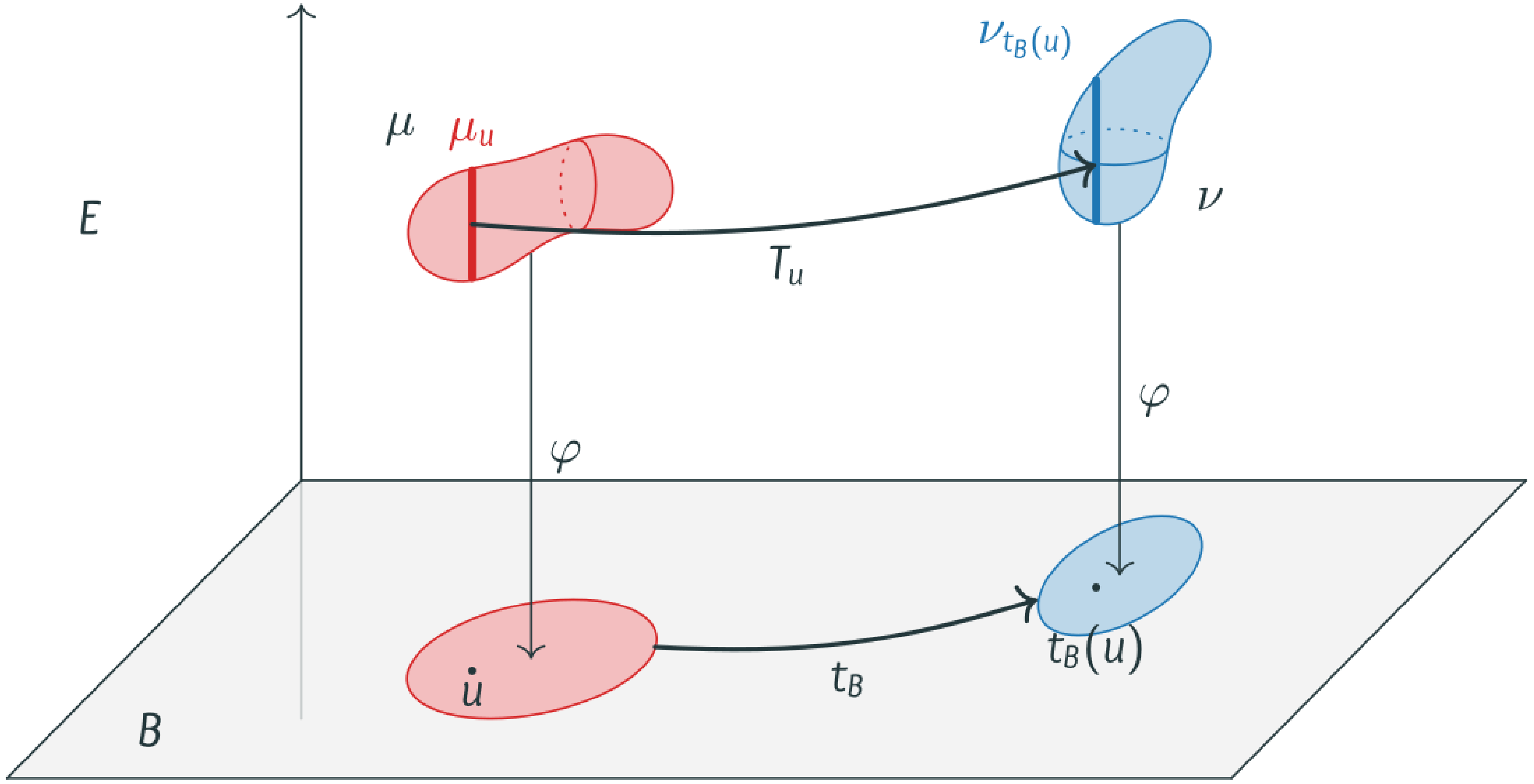
Remark: We have to ensure that the glued map $(u, x) \mapsto T_u(x)$ is measurable.
$\to$ Adapt a result of "measurable selection of transport plans" of Fontbona et al. 2010.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Dumont, L, Vialard, 2024], "Brenier theorem for fiber-invariant costs", Informal

Let $\alpha, \beta \in \mathcal{P}(E)$, κ a cost function, and $\varphi : E \to B$ such that (i) $\varphi\#\mu \ll \mathrm{Vol}(B)$, (ii) $\varphi\#\alpha$-a.e., the fiber $\varphi^{-1}(u)$ is a complete manifold (iii) $\varphi\#\alpha$-a.e., the disintegration α_u of α wrt φ is $\ll \mathrm{Vol}(\varphi^{-1}(u))$, (iv) $\kappa(x,y) = \tilde{\kappa}(\varphi(x),\varphi(y))$ with $\tilde{\kappa}$ twisted on B. Then $\exists$ an optimal transport map between α and β (with some "gradient of convex function" structure on t_B and all $(T_u)_u$).



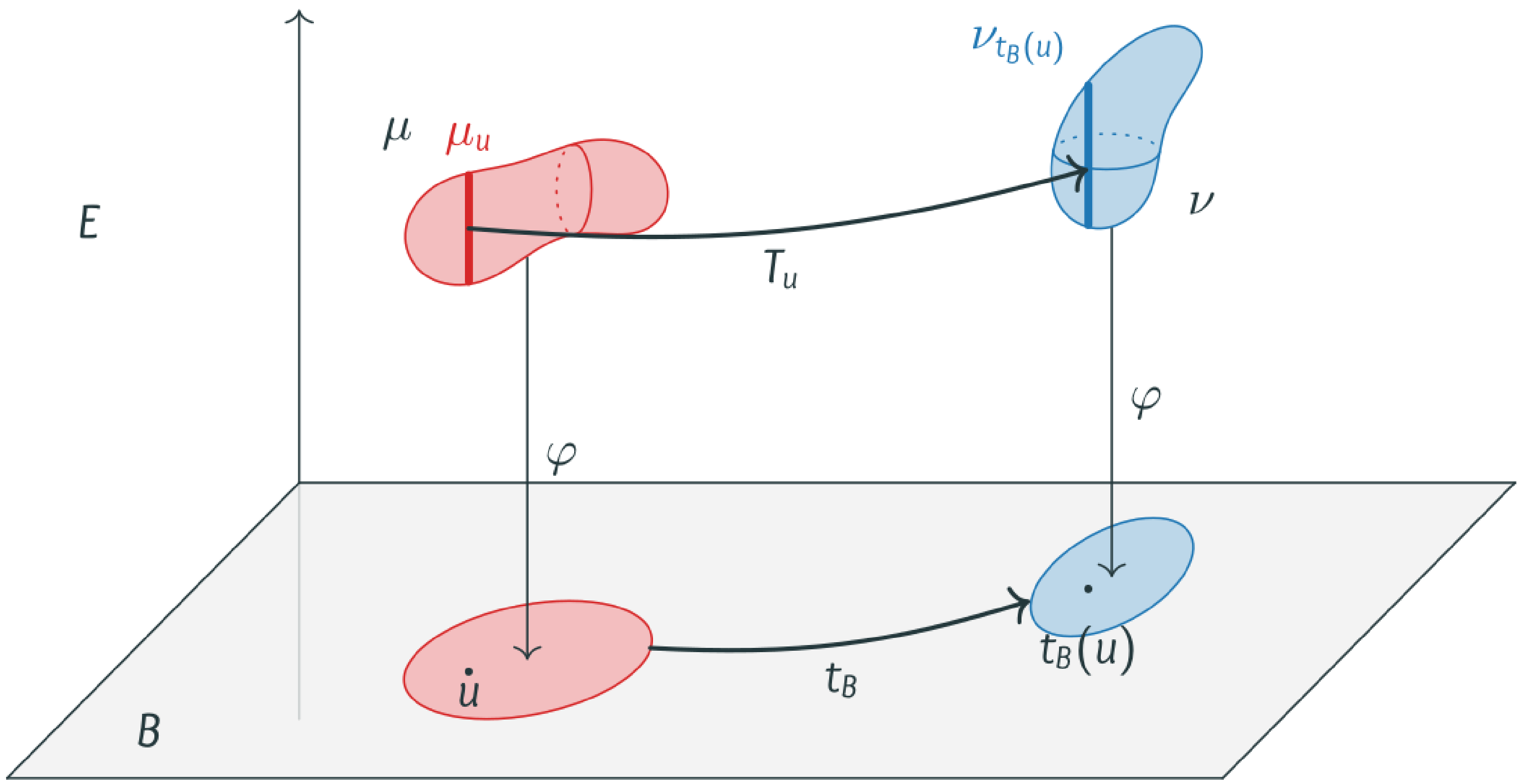
Example: $E = \mathbb{R}^d$, $\kappa(x,y) := (|x| - |y|)^2$, then take $\varphi : x \mapsto |x|$.
Fibers are spheres (except at $|x| = 0$ but we work a.e.), $\tilde{\kappa}$ is the quadratic cost on $B = \mathbb{R} \Rightarrow$ you can prove existence of maps for this cost.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Dumont, L, Vialard, 2024], "Brenier theorem for fiber-invariant costs", Informal

Let $\alpha, \beta \in \mathcal{P}(E)$, κ a cost function, and $\varphi : E \to B$ such that (i) $\varphi\#\mu \ll \mathrm{Vol}(B)$, (ii) $\varphi\#\alpha$-a.e., the fiber $\varphi^{-1}(u)$ is a complete manifold (iii) $\varphi\#\alpha$-a.e., the disintegration α_u of α wrt φ is $\ll \mathrm{Vol}(\varphi^{-1}(u))$, (iv) $\kappa(x,y) = \tilde{\kappa}(\varphi(x),\varphi(y))$ with $\tilde{\kappa}$ twisted on B.
Then $\exists$ an optimal transport map between α and β (with some "gradient of convex function" structure on t_B and all $(T_u)_u$).



Corollary:

There always exist optimal transport maps for the Gromov–Wasserstein problem with inner-product cost.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot,\cdot\rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha,\beta)$. Then, any $\tilde\pi^\star$ minimizes $\mathrm{GW}(\alpha,\beta)$ **if and only if** it minimizes

$$\Pi(\alpha,\beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y')\right)}_{=:C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi\rangle . \tag{OT}$$

Quadratic case: Assume now that $c = |\cdot - \cdot|^2$. We get $C_{\pi^\star}(x,y) = -|x|^2|y|^2 - 4\langle M^\star x, y\rangle$ (still with $M^\star = \int x'^T y' \mathrm{d}\pi^\star(x',y')$).

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x, x') - c(y, y')|^2 \mathrm{d}\pi(x, y)\mathrm{d}\pi(x', y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot, \cdot \rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha, \beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha, \beta)$ **if and only if** it minimizes

$$\Pi(\alpha, \beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x, x') - c(y, y')|^2 \mathrm{d}\pi^\star(x', y') \right)}_{=: C_{\pi^\star}(x,y)} \mathrm{d}\pi(x, y) = \langle C_{\pi^\star}, \pi \rangle . \quad \text{(OT)}$$

Quadratic case: Assume now that $c = |\cdot - \cdot|^2$. We get $C_{\pi^\star}(x, y) = -|x|^2|y|^2 - 4 \langle M^\star x, y \rangle$ (still with $M^\star = \int x'^T y' \mathrm{d}\pi^\star(x', y')$).

Proposition:

If $M^\star$ is of rank d or $d-1$, the 2-twist condition is satisfied (idea: solutions of the equation $\nabla_x c(x, y) = \nabla_x c(x, y')$ are given by the intersection of a 1D-line and a $d-1$-sphere), $\Rightarrow$ there exist "2-maps".
If $M^\star$ is of rank $\leqslant d - 2$...

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot,\cdot \rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha,\beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha,\beta)$ **if and only if** it minimizes

$$\Pi(\alpha,\beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y') \right)}_{=: C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi \rangle \,. \tag{OT}$$

Quadratic case: Assume now that $c = |\cdot - \cdot|^2$. We get $C_{\pi^\star}(x,y) = -|x|^2|y|^2 - 4\langle M^\star x, y\rangle$ (still with $M^\star = \int x'^T y' \mathrm{d}\pi^\star(x',y')$).

Proposition:

If $M^\star$ is of rank d or $d-1$, the 2-twist condition is satisfied (idea: solutions of the equation $\nabla_x c(x,y) = \nabla_x c(x,y')$ are given by the intersection of a 1D-line and a $d-1$-sphere), $\Rightarrow$ there exist "2-maps".
If $M^\star$ is of rank $\leqslant d-2$... there exists an optimal transport map!

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot, \cdot \rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha, \beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha, \beta)$ **if and only if** it minimizes

$$\Pi(\alpha, \beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y') \right)}_{=: C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi \rangle \,. \quad \text{(OT)}$$

Quadratic case: Assume now that $c = |\cdot - \cdot|^2$. We get $C_{\pi^\star}(x,y) = -|x|^2|y|^2 - 4 \langle M^\star x, y \rangle$ (still with $M^\star = \int x'^T y' \mathrm{d}\pi^\star(x',y')$).

Proposition:

If $M^\star$ is of rank d or $d-1$, the 2-twist condition is satisfied (idea: solutions of the equation $\nabla_x c(x,y) = \nabla_x c(x,y')$ are given by the intersection of a 1D-line and a $d-1$-sphere), $\Rightarrow$ there exist "2-maps".
If $M^\star$ is of rank $\leqslant d-2$... there exists an optimal transport map!

Sketch of proof: This time, the "projection" φ is of the form $\varphi(x) = (x_H, |x_\perp|^2)$, with $x_H \in \mathbb{R}^{\mathrm{rk}(M^\star)}$ and $x_\perp \in \mathbb{R}^{d - \mathrm{rk}(M^\star)}$, the resulting cost $\tilde{\kappa}$ *is twisted* and the fibers are spheres of dimension $d - 1 - \mathrm{rk}(M^\star)$.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Proposition: [Konno 1976, Séjourné et al. 2021]

Let $c = |\cdot - \cdot|^2$ or $c = \langle \cdot, \cdot \rangle$. Let $\pi^\star$ solve $\mathrm{GW}(\alpha,\beta)$. Then, any $\tilde{\pi}^\star$ minimizes $\mathrm{GW}(\alpha,\beta)$ **if and only if** it minimizes

$$\Pi(\alpha,\beta) \ni \pi \mapsto \int_{x,y} \underbrace{\left(\int_{x',y'} |c(x,x') - c(y,y')|^2 \mathrm{d}\pi^\star(x',y') \right)}_{=:C_{\pi^\star}(x,y)} \mathrm{d}\pi(x,y) = \langle C_{\pi^\star}, \pi \rangle . \quad \text{(OT)}$$

Quadratic case: Assume now that $c = |\cdot - \cdot|^2$. We get $C_{\pi^\star}(x,y) = -|x|^2|y|^2 - 4\langle M^\star x, y\rangle$ (still with $M^\star = \int x'^T y' \mathrm{d}\pi^\star(x',y')$).

Proposition:

If $M^\star$ is of rank d or $d-1$, the 2-twist condition is satisfied (idea: solutions of the equation $\nabla_x c(x,y) = \nabla_x c(x,y')$ are given by the intersection of a 1D-line and a $d-1$-sphere), $\Rightarrow$ there exist "2-maps".
If $M^\star$ is of rank $\leqslant d-2$... there exists an optimal transport map!

Sketch of proof: This time, the "projection" φ is of the form $\varphi(x) = (x_H, |x_\perp|^2)$, with $x_H \in \mathbb{R}^{\mathrm{rk}(M^\star)}$ and $x_\perp \in \mathbb{R}^{d-\mathrm{rk}(M^\star)}$, the resulting cost $\tilde{\kappa}$ *is twisted* and the fibers are spheres of dimension $d - 1 - \mathrm{rk}(M^\star)$.

Remark: This argument fails when $\mathrm{rk}(M^\star) = d-1$ because fibers are $\{\pm x_\perp\}$ $\rightarrow$ cannot build maps between fibers.

Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x, x') - c(y, y')|^2 \mathrm{d}\pi(x, y) \mathrm{d}\pi(x', y')$.

Tightness? For the quadratic cost $c = |\cdot - \cdot|^2$, we proved that if $\mathrm{rk}(M^\star) \leqslant d - 2$, there exists an optimal Monge map, but we only have 2-maps if $M^\star$ is of higher rank... Did we miss something?

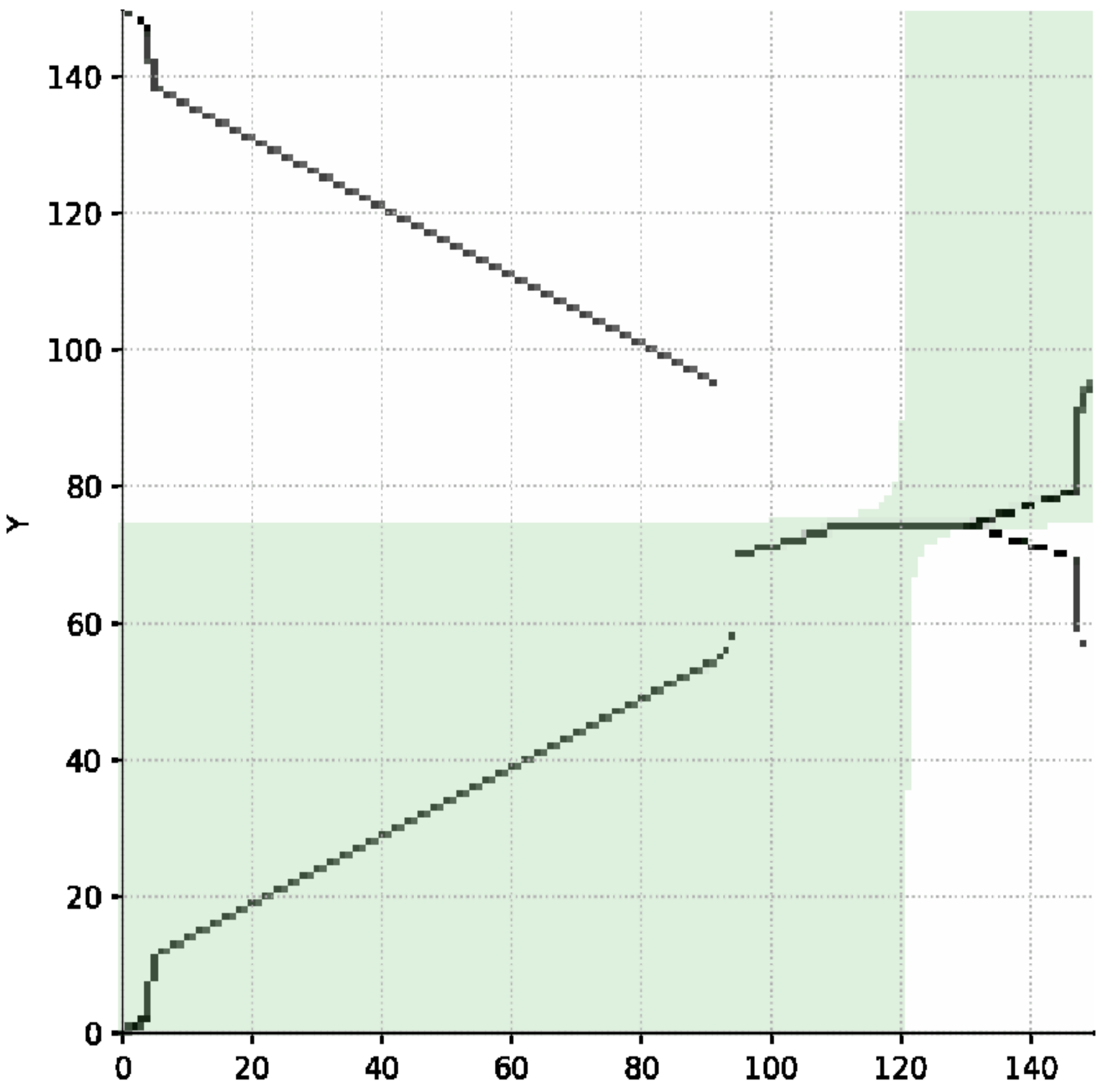
Optimal transport maps for GW

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x, x') - c(y, y')|^2 \mathrm{d}\pi(x, y)\mathrm{d}\pi(x', y')$.

Tightness? For the quadratic cost $c = |\cdot - \cdot|^2$, we proved that if $\mathrm{rk}(M^\star) \leqslant d-2$, there exists an optimal Monge map, but we only have 2-maps if $M^\star$ is of higher rank... Did we miss something?

Conjecture: No (only numerical evidence though).

More precisely, we adversarially build configurations for which the optimal GW plan looks like an actual 2-map.



Remark: Several layers of approximation/discretization :

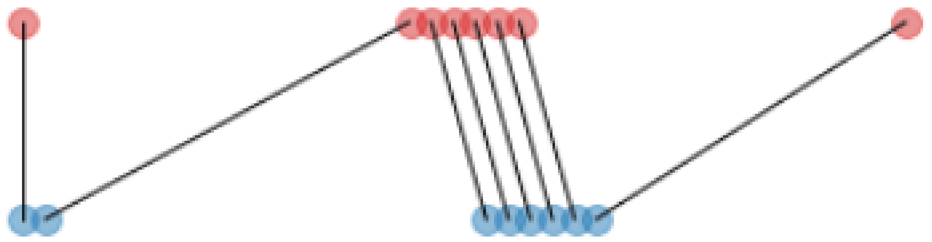
- Discretize the ground space $\mathbb{R}$ (starting from α with density),
- in dim 1, $M^\star \in [m_{\min}, m_{\max}]$ with tractable $m_{\min/\max}$, and thus solve the OT problem with cost $c_m(x, y) = -x^2y^2 - 4mxy$ for m in a discretization of this segment, yielding $\pi^\star_m$, and then evaluate the GW performance of all $\pi^\star_m$ to find the actual optimal GW plan $\to$ need stability results to make this rigorous.

GW IN DIMENSION 1

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Observation: Assume $d = 1$, $c = |\cdot - \cdot|^2$, and let $\alpha = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\beta = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, with $x_1 \leqslant \ldots \leqslant x_n$ and $y_1 \leqslant \ldots y_n$.
It would be nice if the increasing mapping $(x_i \mapsto y_i)$ or the decreasing one $(x_i \mapsto y_{n+1-i})$ would be optimal.

- Empirically satisfied, works well in applications...



$\sigma = \mathrm{id}$

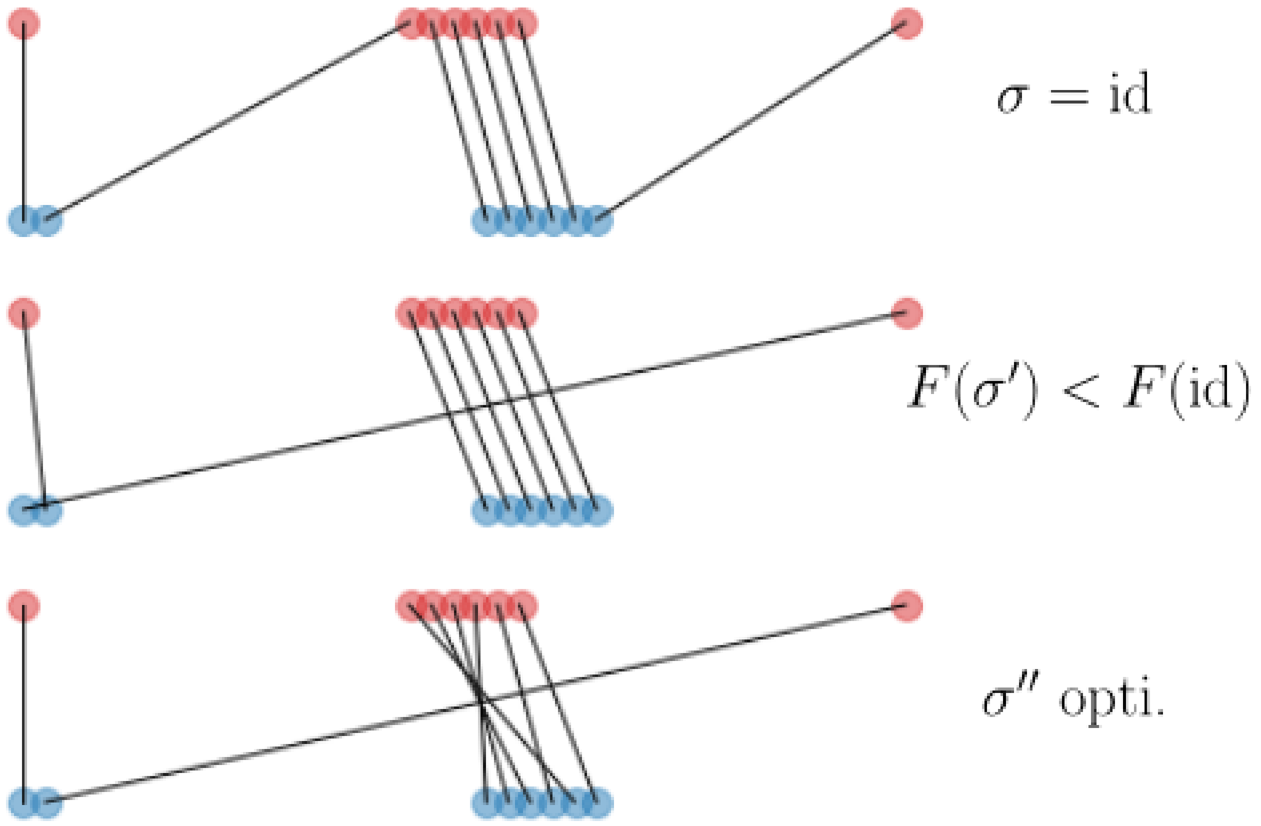
GW in dimension 1

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Observation: Assume $d = 1$, $c = |\cdot - \cdot|^2$, and let $\alpha = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\beta = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, with $x_1 \leqslant \ldots \leqslant x_n$ and $y_1 \leqslant \ldots y_n$.
It would be nice if the increasing mapping $(x_i \mapsto y_i)$ or the decreasing one $(x_i \mapsto y_{n+1-i})$ would be optimal.

- Empirically satisfied, works well in applications...

... But [Beinhart et al., 2022] provides a counter-example!



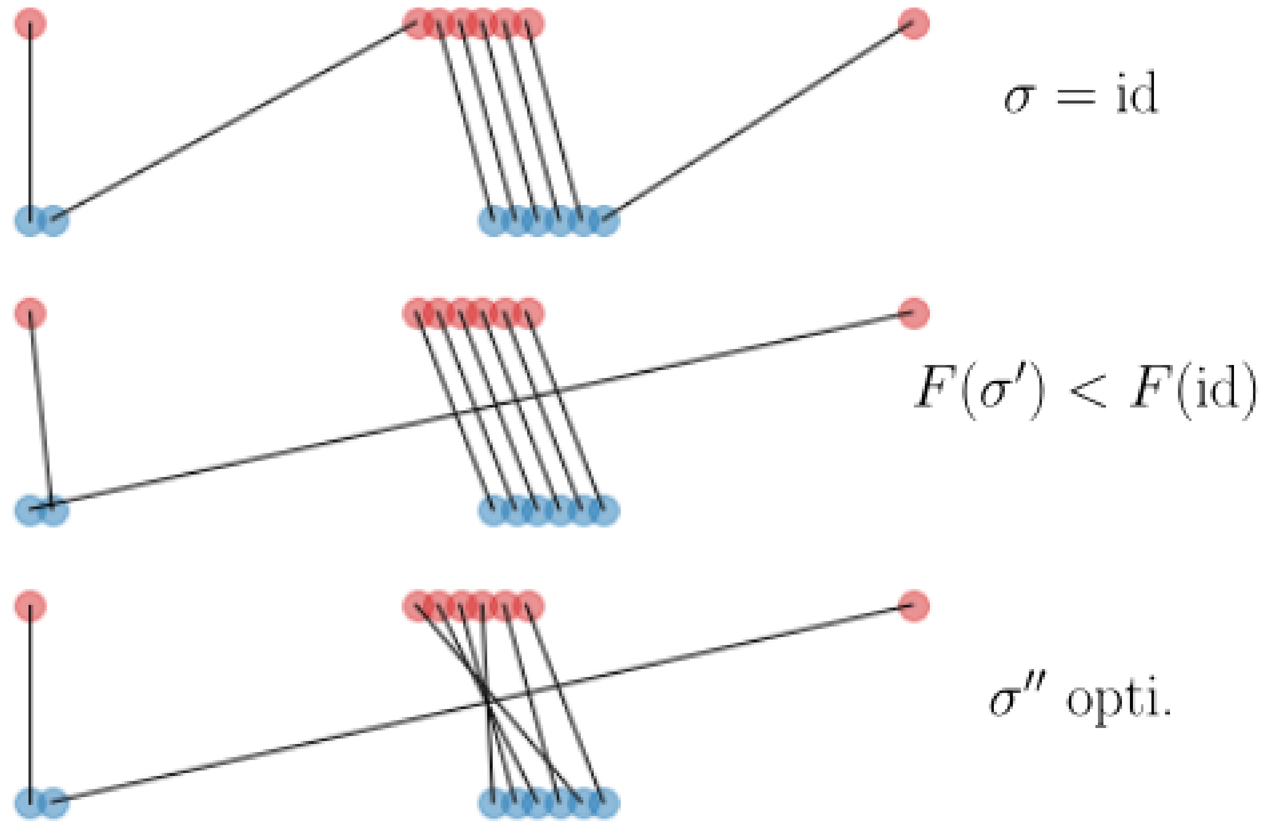
GW IN DIMENSION 1

Recall: $\mathrm{GW}(\alpha,\beta) = \inf_{\pi\in\Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Observation: Assume $d=1$, $c = |\cdot - \cdot|^2$, and let $\alpha = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\beta = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, with $x_1 \leqslant \ldots \leqslant x_n$ and $y_1 \leqslant \ldots y_n$.
It would be nice if the increasing mapping ($x_i \mapsto y_i$) or the decreasing one ($x_i \mapsto y_{n+1-i}$) would be optimal.

- Empirically satisfied, works well in applications...

... But [Beinhart et al., 2022] provides a counter-example!



(Open) Question: Why does it work "in practice" and only fails on "adversarial" examples?

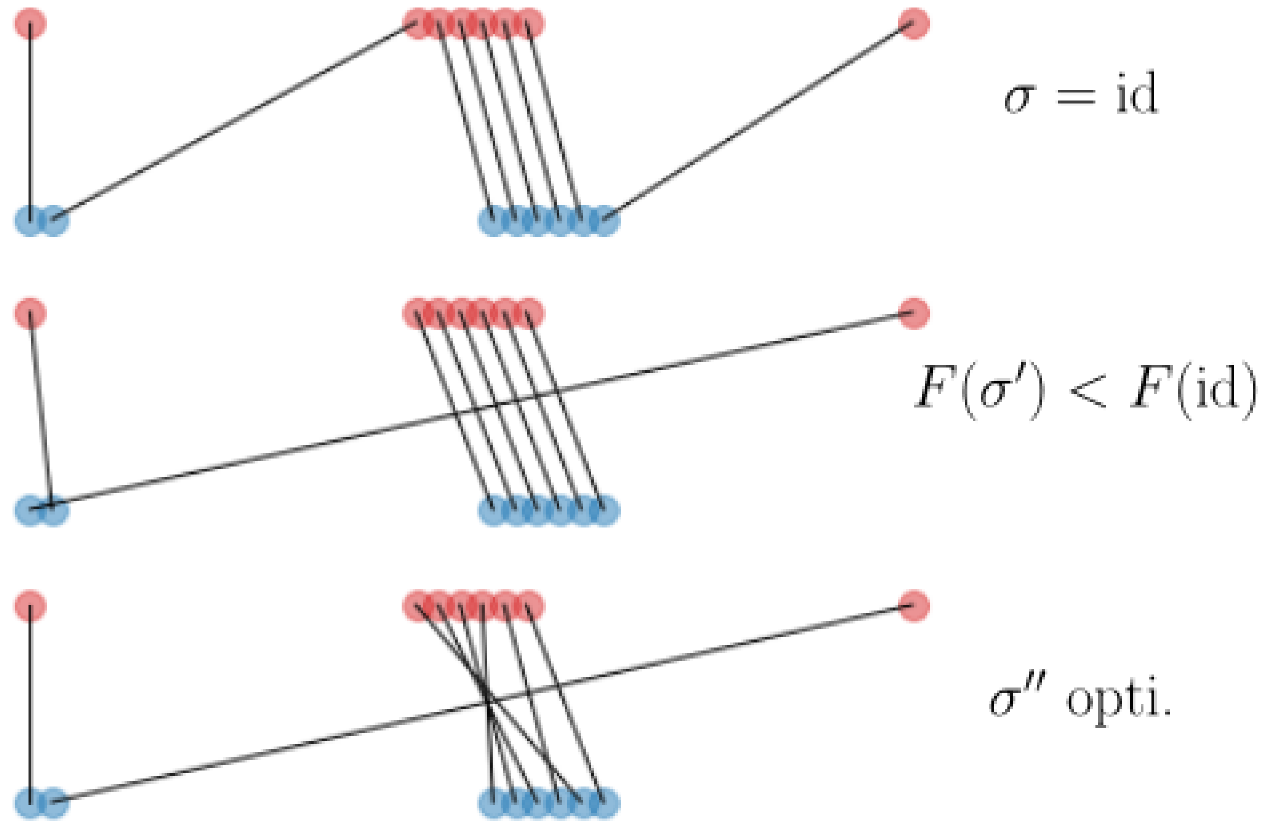
GW in dimension 1

Recall: $\mathrm{GW}(\alpha, \beta) = \inf_{\pi \in \Pi(\alpha,\beta)} \iint |c(x,x') - c(y,y')|^2 \mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')$.

Observation: Assume $d = 1$, $c = |\cdot - \cdot|^2$, and let $\alpha = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\beta = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, with $x_1 \leqslant \ldots \leqslant x_n$ and $y_1 \leqslant \ldots y_n$.
It would be nice if the increasing mapping $(x_i \mapsto y_i)$ or the decreasing one $(x_i \mapsto y_{n+1-i})$ would be optimal.

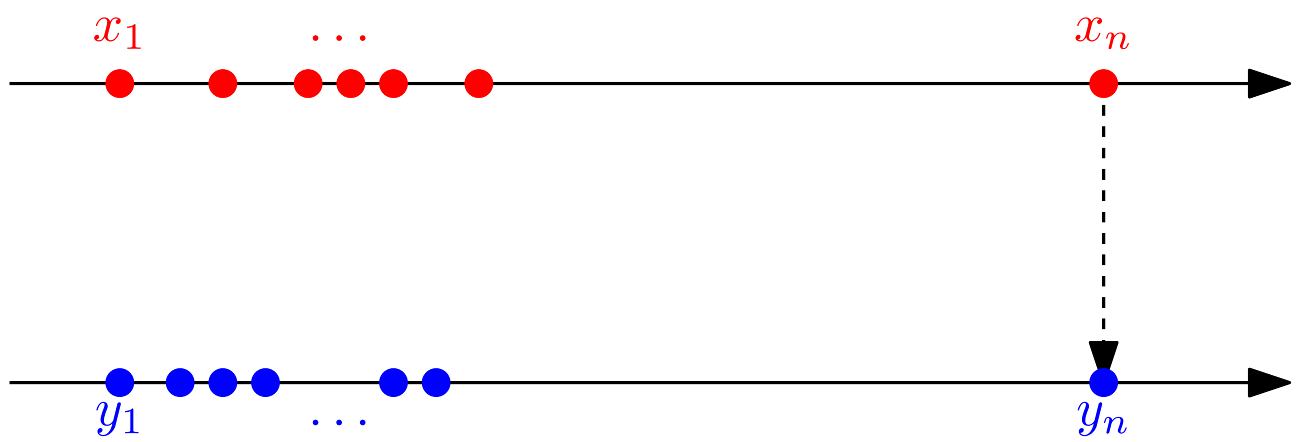
- Empirically satisfied, works well in applications...

... But [Beinhart et al., 2022] provides a counter-example!



(Open) Question: Why does it work "in practice" and only fails on "adversarial" examples?

(Partial) Answer (informal): if (say) x_n and y_n are far away from the others $(x_i, y_i)_i$, they must be matched together, and this "forces" a monotone (increasing) matching ("long-range interactions dominate").



Conclusion & Some remaining questions

Conclusion :

- $\exists$ (Gromov–)Monge map for GW with the inner-product cost,
- $\exists$ 2-maps for GW with quadratic cost, and maps provided the covariance of a solution $\pi^\star$ is sufficiently singular. Numerical examples suggests the tightness of this result.
- (Gromov–)Monge maps are obtained by gluing a (standard) Monge map on a projection space B and Monge maps fiber-wise.
- In 1D (and probably in higher-dimension), long-range interactions dominate and dictate the structure of the matching.

Some questions left:

- Is $M^\star$ generically non-singular?
- For the quadratic cost: can we find a condition to guarantee the existence of Monge map even if $M^\star$ is of full rank? (Finding a numerical counter-example was actually quite hard!)
- Can we leverage more regularity on the disintegration $(\mu_u)_u, (\nu_u)_u$ to get more regularity/structure on the Gromov–Monge map?
- Can we leverage the existence of Gromov–Monge maps in practice?
- In 1D, can we make our observations more quantitative?

Thanks :-)