Barycenters for transport costs

Julie Delon, Nathaël Gozlan and Eloi Tanguy

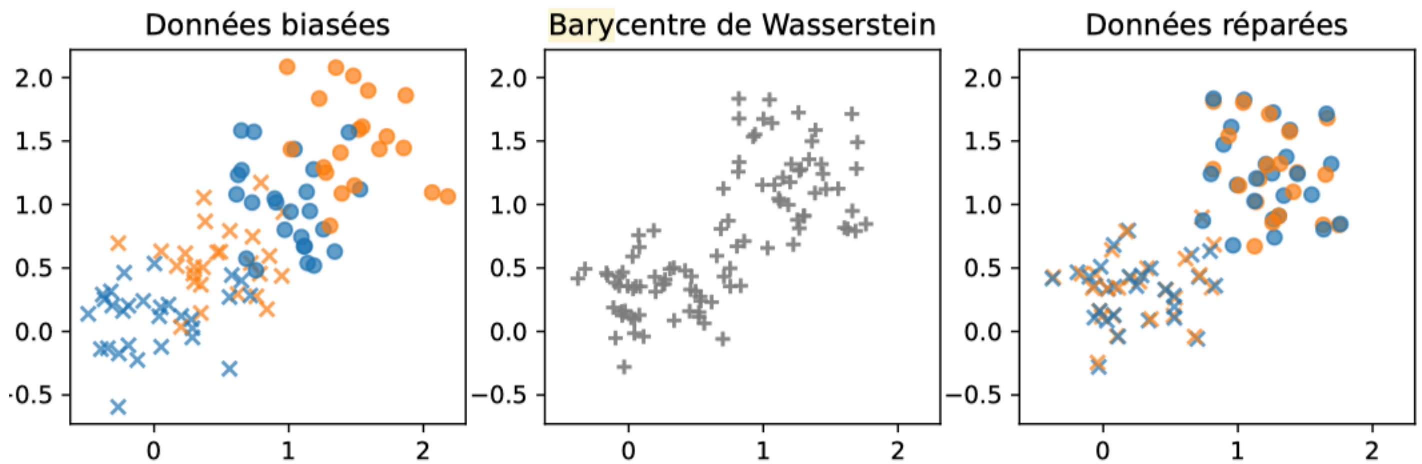
LABORATOIRE MAP5

Université Paris Cité

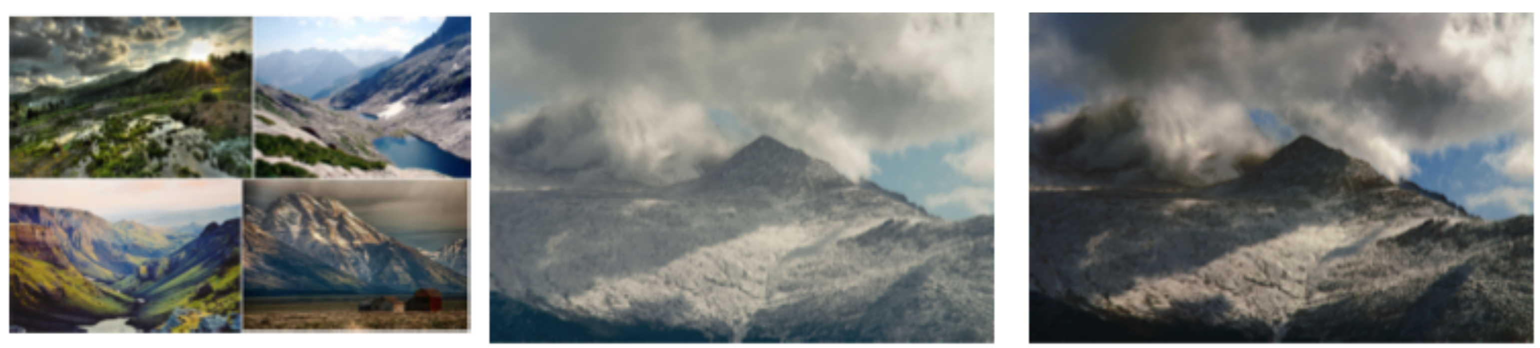
Barycenters in data science

[Simon et al.,2019]

[Levy, 2015]

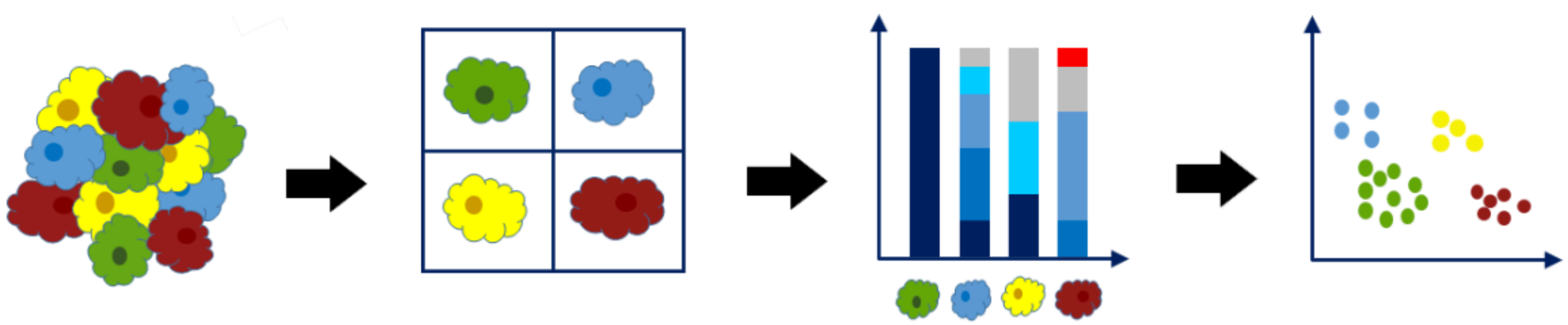


OT barycenters for Fairness [Gordaliza et al., 2019]



Color harmonization [Bonneel et al. 2016]

Genomics [Schiebinger et al. 2019]

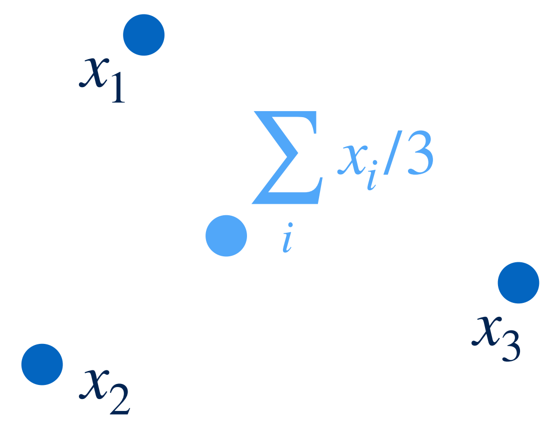


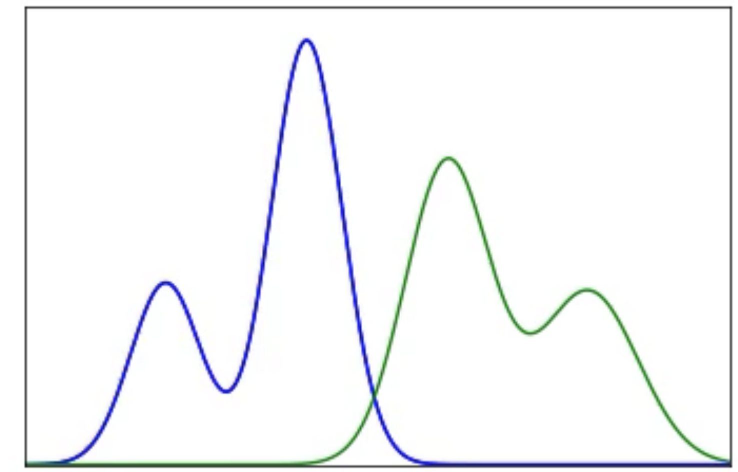
[Vacher et al., 2020]

Barycenters

- $(x_1, \ldots, x_p)$ **points** of $\mathscr{X}$
- $(\lambda_1, \ldots, \lambda_p)$ **weights**, ≥ 0

  and s.t. $\sum_i \lambda_i = 1$
- **costs** $c_i : \mathscr{X} \times \mathscr{X} \to \mathbb{R}^+$





$$B(x_1, \ldots, x_p) \in \underset{x}{\operatorname{argmin}}\ \lambda_1 c_1(x, x_1) + \ldots + \lambda_p c_p(x, x_p)$$

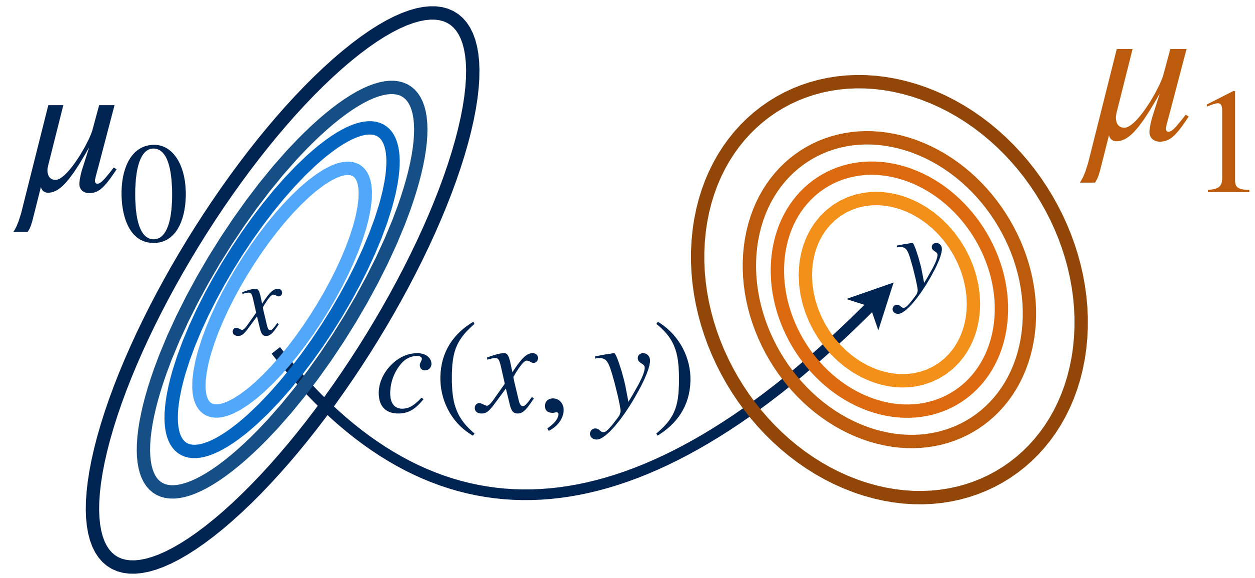
<u>Ex</u> : $\mathscr{X} = \mathbb{R}^d,\ c_i = \|.\|^2$ ⟶ weighted mean

$\mathscr{X} = \mathbb{R}^d,\ c_i = \|.\|$ ⟶ weighted geometric median

$\mathscr{X} = \mathbb{R}^d,\ c_i = \|.\|^p$ $\quad \ldots$

Transport costs and Wasserstein distances

- **cost** $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$
- $\Pi(\mu_0, \mu_1)$ = couplings between μ_0 and μ_1



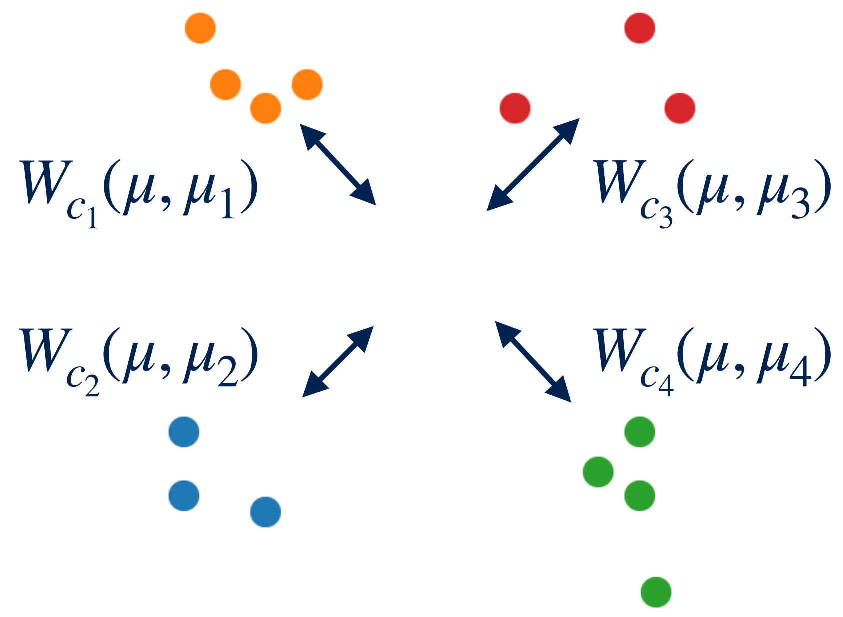
$$W_c(\mu_0, \mu_1) = \inf_{\gamma \in \Pi(\mu_0, \mu_1)} \iint c(x, y) d\gamma(x, y)$$

Theorem: If $c(x, y) = d(x, y)^p$ with $p \geq 1$ and d a distance on $\mathcal{X}$, $W_p(\mu_0, \mu_1) := W_c(\mu_0, \mu_1)^{\frac{1}{p}}$ defines a **distance between probability measures**.

p=2 or 1 used in most applications

Fréchet means / Wasserstein barycenters

- $(\mu_1, \ldots, \mu_p)$ **proba. measures**
- $(\lambda_1, \ldots, \lambda_p)$ **weights**, ≥ 0
  and s.t. $\sum_i \lambda_i = 1$



$$B(\mu_1, \ldots, \mu_p) \in \operatorname*{argmin}_{\rho} \lambda_1 W_{c_1}(\rho, \mu_1) + \ldots + \lambda_p W_{c_p}(\rho, \mu_p)$$

- [Carlier Ekeland, 2010] → continuous ground costs
- [Agueh, Carlier 2011] → W_2^2, existence + uniqueness if the μ_i are AC.
- [Carlier et al. 2023] → W_1, Wasserstein medians
- [Legouic et al. 2016], [Brizzi et al., 2024] → W_p for $p > 1$, W_h for specific h.

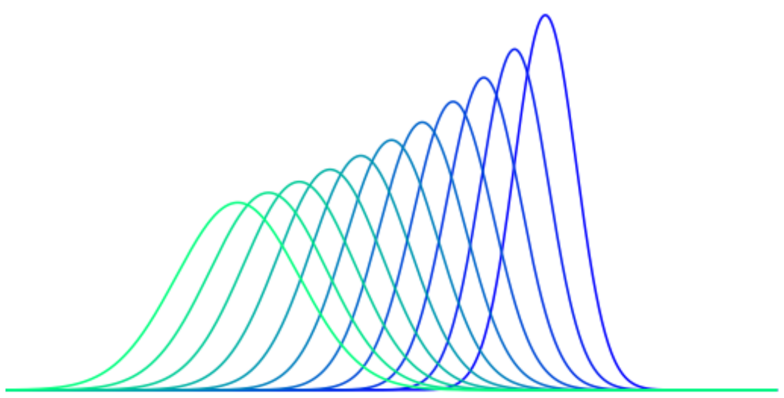
Barycenter between Gaussians for W_2

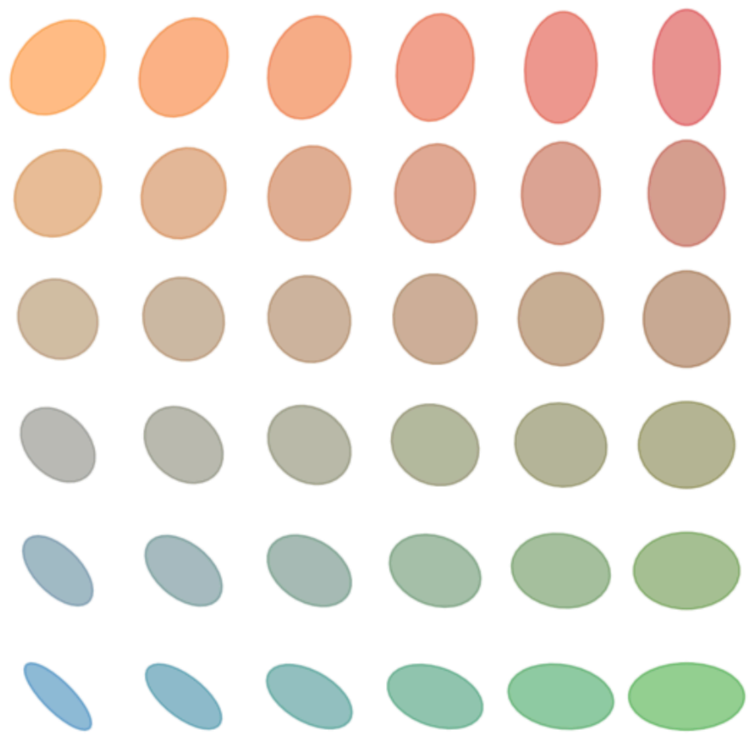
$\mu_i = \mathcal{N}(m_i, \Sigma_i), i \in \{1, \dots, p\}$ Gaussian distributions on $\mathbb{R}^d$

Barycenter [Agueh, Carlier 2011]:

$$\mathcal{N}(m^*, \Sigma^*) = \underset{\rho}{\operatorname{argmin}} \sum_{i=1}^{p} \lambda_i W_2^2(\mu_i, \rho)$$

$$m^* = \sum_{i=1}^{p} \lambda_i m_i, \ \Sigma^* = \min_{\Sigma} \sum_{i=1}^{p} \lambda_i \mathrm{B}^2(\Sigma, \Sigma_i)$$





$$B^2(\Sigma_0, \Sigma_1) = \operatorname{tr}\left(\Sigma_0 + \Sigma_1 - 2\left(\Sigma_0^{\frac{1}{2}} \Sigma_1 \Sigma_0^{\frac{1}{2}}\right)^{\frac{1}{2}}\right)$$

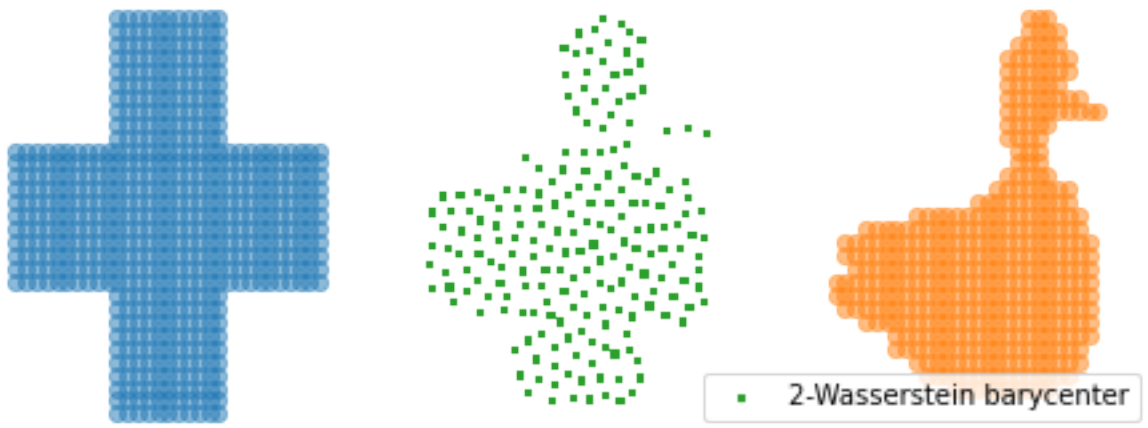
Computing barycenters for transport costs

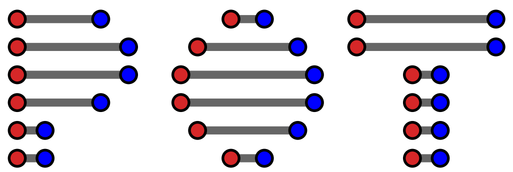
NP-hard problem [Altschuler et al. 21]

Exponential dependence on d or p

Some Algorithms

- Linear Programming-MMOT
- Sliced/Radon OT [Rabin et al.12, .15].
- Entropic barycenters [Benamou et al 15 Solomon et al 15,...]
- Free support [Cuturi, Doucet 14]





Free support barycenter: a fixed point problem

- $(\mu_1, \dots, \mu_p)$ **proba. measures**
- $(\lambda_1, \dots, \lambda_p)$ **weights**, ≥ 0

  and s.t. $\sum_i \lambda_i = 1$

Prop **[Alvarez-Esteban et al., 15]**

- W_2 + AC measures $\Rightarrow$barycenters $\in$ fixed points of G.
- **CV** of subsequence to fixed point.
- If unique fp, it is a barycenter

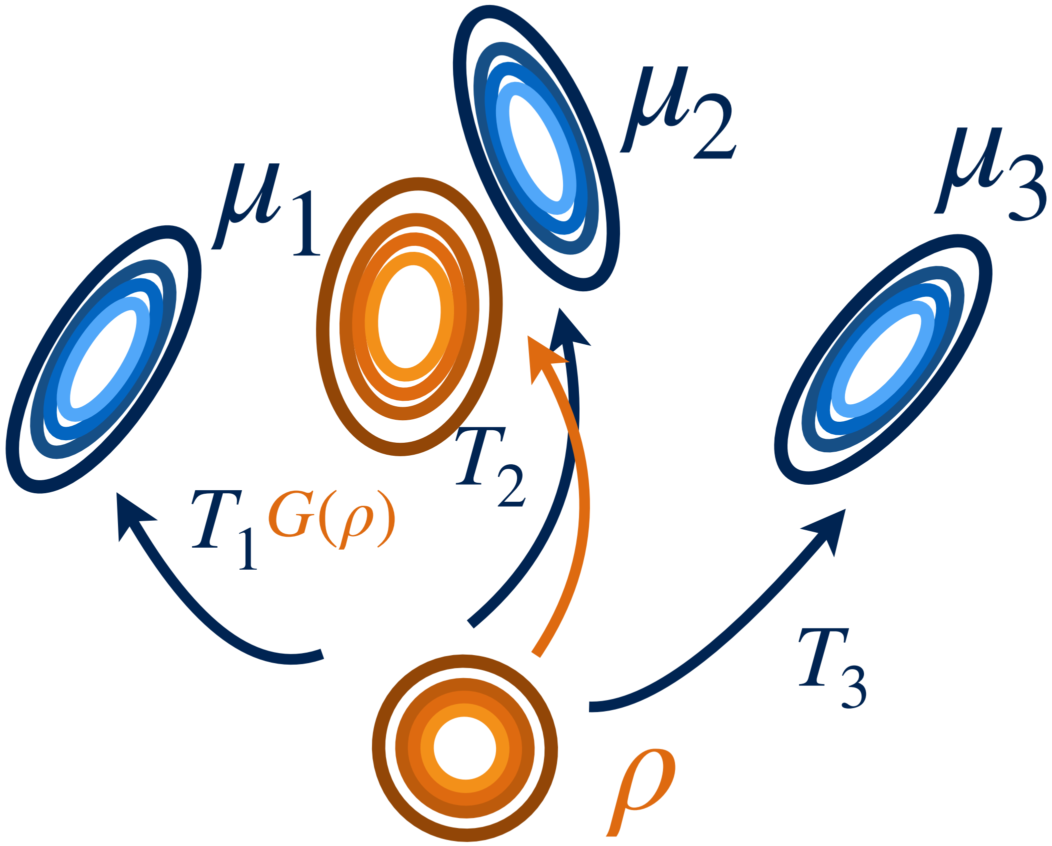
ALGO:

$$\rho_{n+1} = G(\rho_n)$$

Def: $X \sim \rho$

T_k OT maps between ρ and μ_k

$$G(\rho) = \mathscr{L}\left(\sum \lambda_j T_j(X)\right)$$



Beyond W_p: generalized barycenters

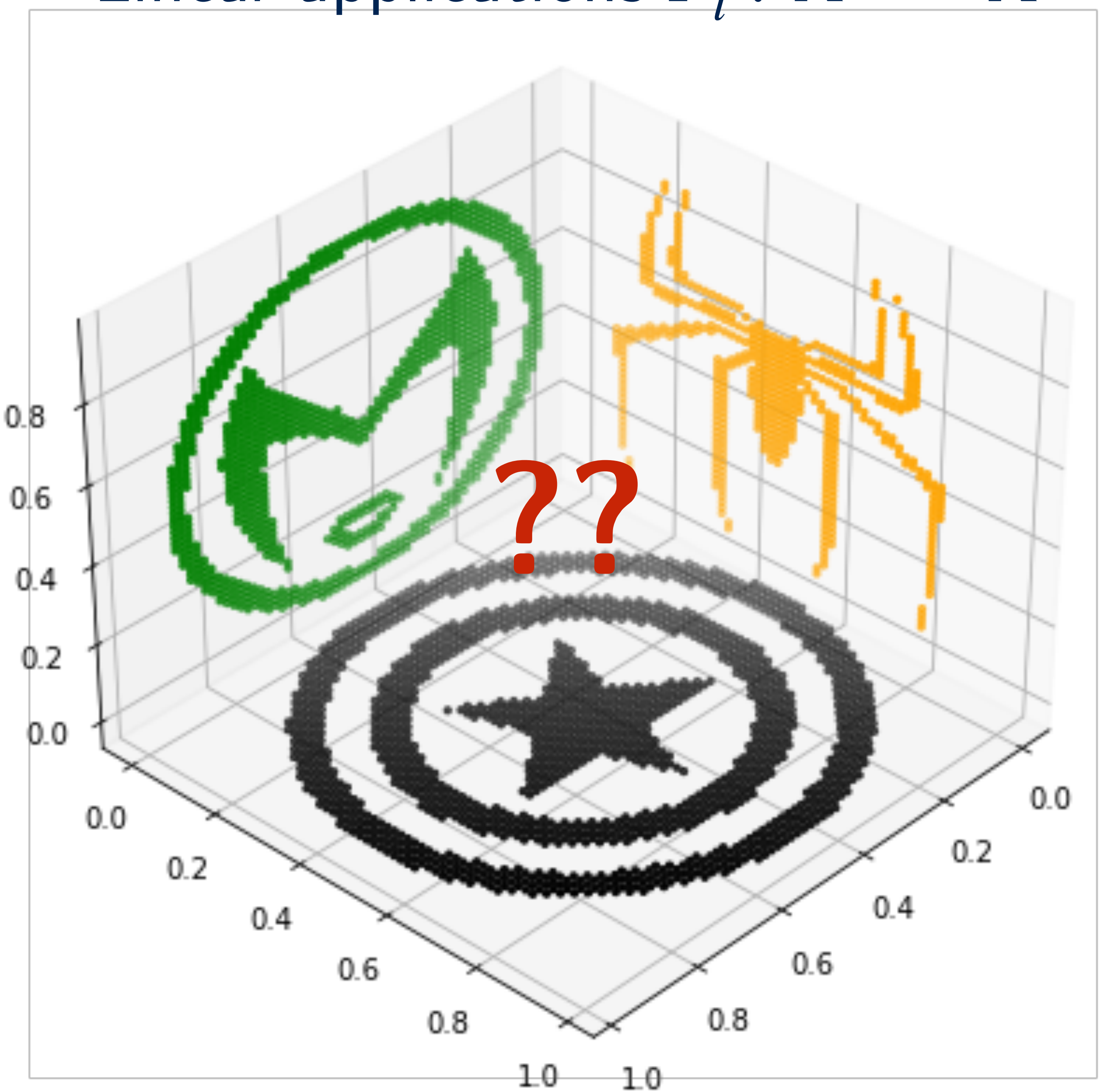
- $(\mu_1, \ldots, \mu_p)$ **proba. measures on** $\mathbb{R}^{d_i}$
- $(\lambda_1, \ldots, \lambda_p)$ **weights**, ≥ 0

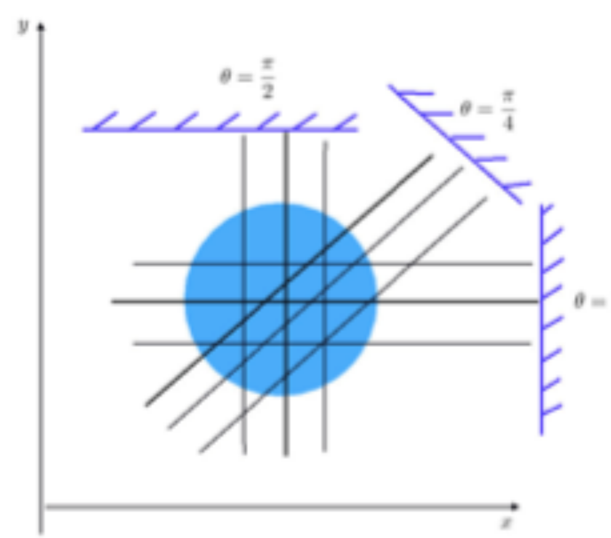
  and s.t. $\sum \lambda_i = 1$
- Linear applications $P_i : \mathbb{R}^d \to \mathbb{R}^{d_i}$

Generalized barycenters

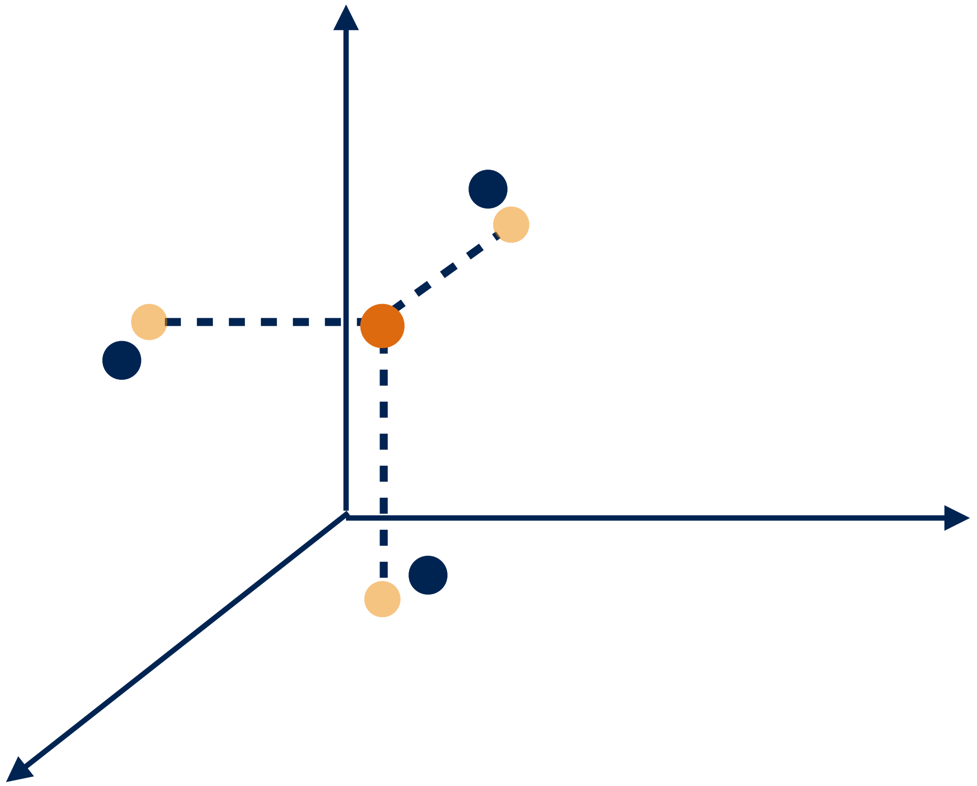
[D. et al, 2022]

$$\inf_{\gamma \in \mathscr{P}_2(\mathbb{R}^d)} \underbrace{\sum_{i=1}^{p} \lambda_i W_2^2(\mu_i, P_i \# \gamma)}_{\mathscr{F}(\gamma)}$$





Solutions for generalized barycenters?



$$B(x_1, \ldots, x_p) \in \text{argmin}_{x \in \mathbb{R}^d} \sum_{i=1}^{p} \lambda_i \|P_i x - x_i\|^2$$

Hyp: $A := \sum_{i=1}^{p} \lambda_i P_i^T P_i$ full rank

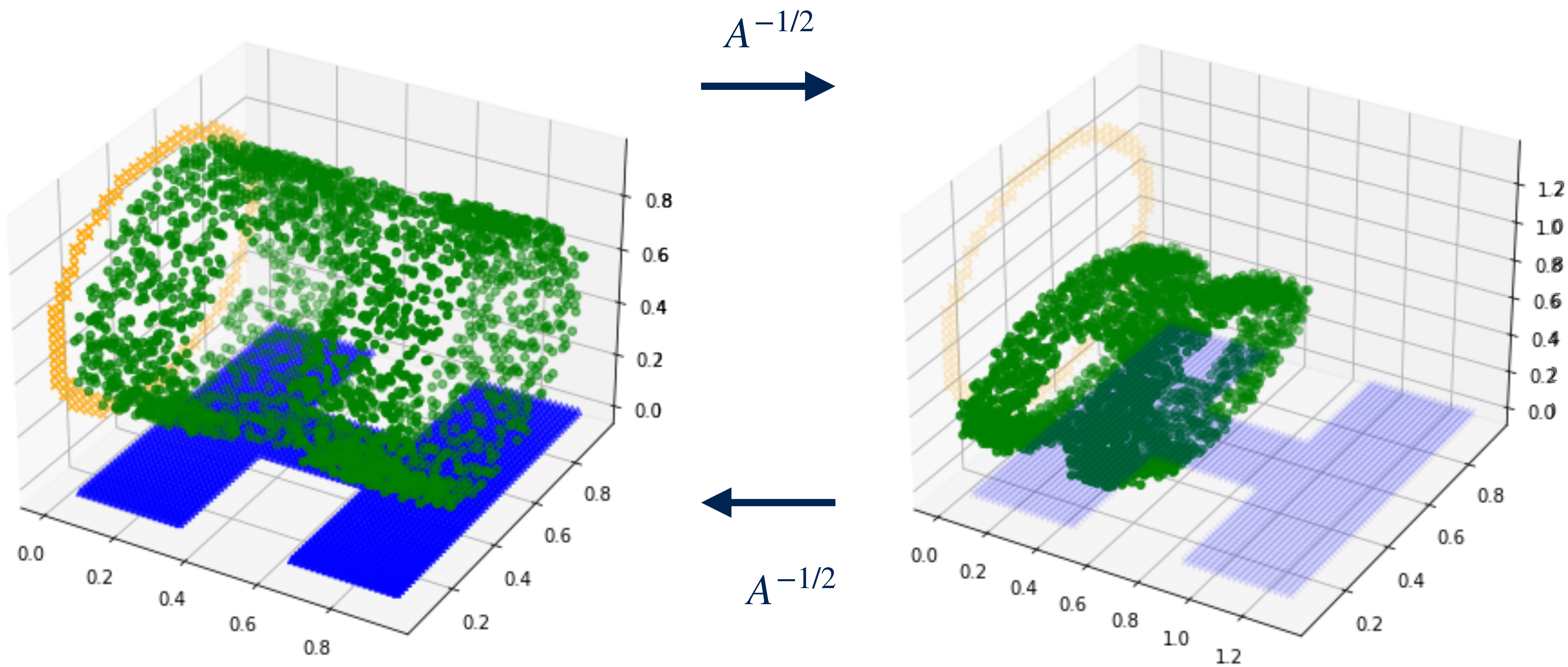
Solution: $\hat{x} = A^{-1}(\sum_{i=1}^{p} \lambda_i P_i^T x_i)$

Generalized barycenters

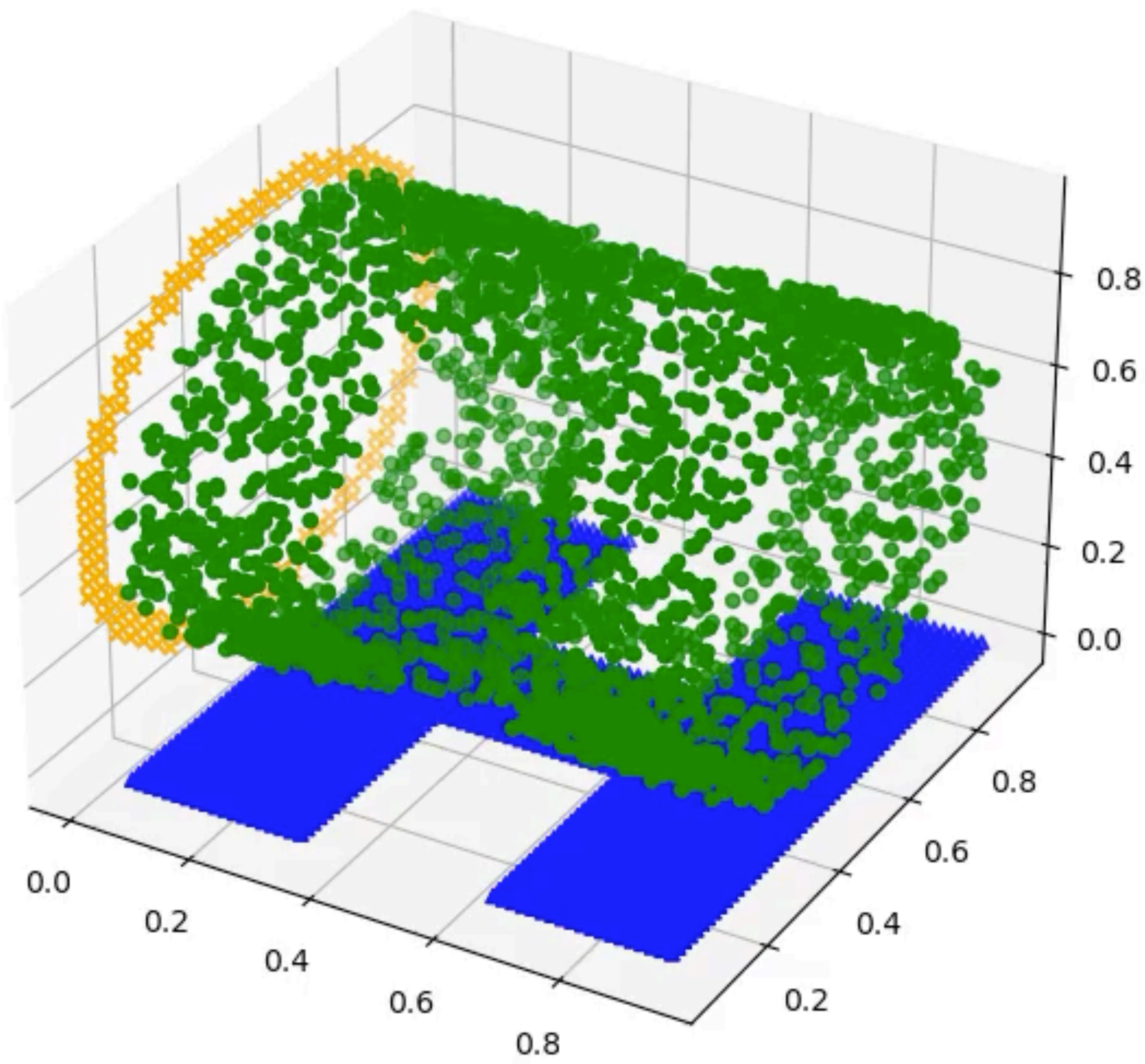
γ^* solution **iff** $A^{1/2} \# \gamma^*$ is a W_2^2-barycenter of the measures $(A^{-1/2} P_i^T) \# \mu_i$

Computing generalized barycenters

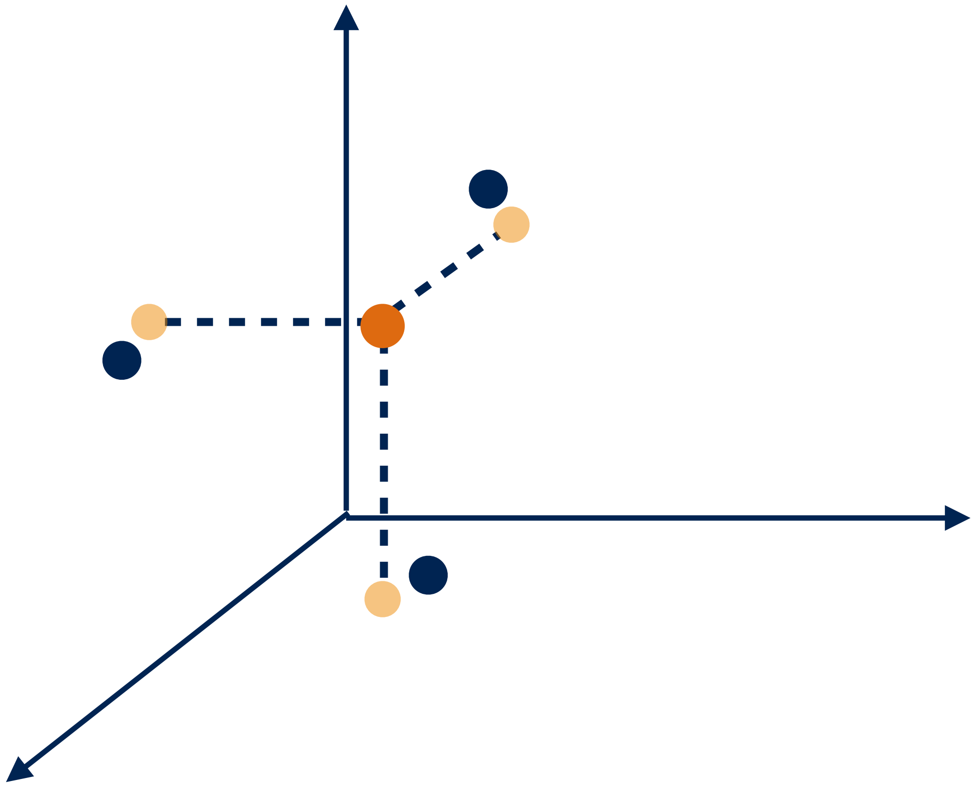
2 distributions



Computing generalized barycenters



Solutions for generalized barycenters?



$$B(x_1, \ldots, x_p) \in \operatorname{argmin}_{x \in \mathbb{R}^d} \sum_{i=1}^{p} \lambda_i \|P_i x - x_i\|^2$$

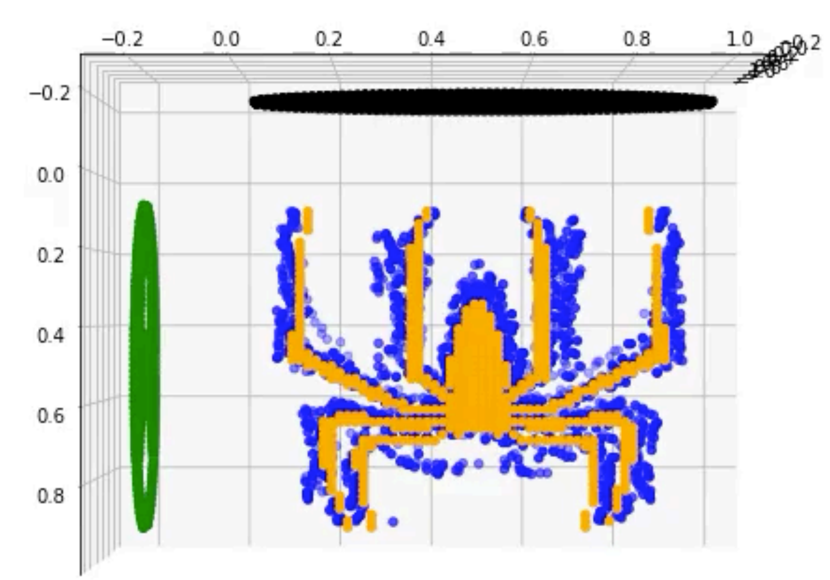
Hyp: $A := \sum_{i=1}^{p} \lambda_i P_i^T P_i$ full rank

Solution: $\hat{x} = A^{-1}(\sum_{i=1}^{p} \lambda_i P_i^T x_i)$

<u>**Generalized barycenters**</u>

γ^* solution **iff** $A^{1/2} \# \gamma^*$ is a W_2^2-barycenter of the measures $(A^{-1/2} P_i^T) \# \mu_i$

Existence, no uniqueness
Gaussian case
Link with multi marginal transport



Included in POT by Eloi Tanguy

What about generic ground costs?

$$\underset{\rho}{\operatorname{argmin}}\ \lambda_1 W_{c_1}(\rho, \mu_1) + \ldots + \lambda_p W_{c_p}(\rho, \mu_p)$$

$\rho_{n+1} = G(\rho_n)$??

Hyp:

- continuous costs c_k
- $B(x_1, \ldots, x_p) = \underset{x}{\operatorname{argmin}}\ \lambda_1 c_1(x, x_1) + \ldots + \lambda_p c_p(x, x_p)$ (unique)

Def $X \sim \rho$

T_k OT maps between ρ and μ_k

$G(\rho) = \mathscr{L}(B(T_1(X), \ldots, T_p(X)))$

↓

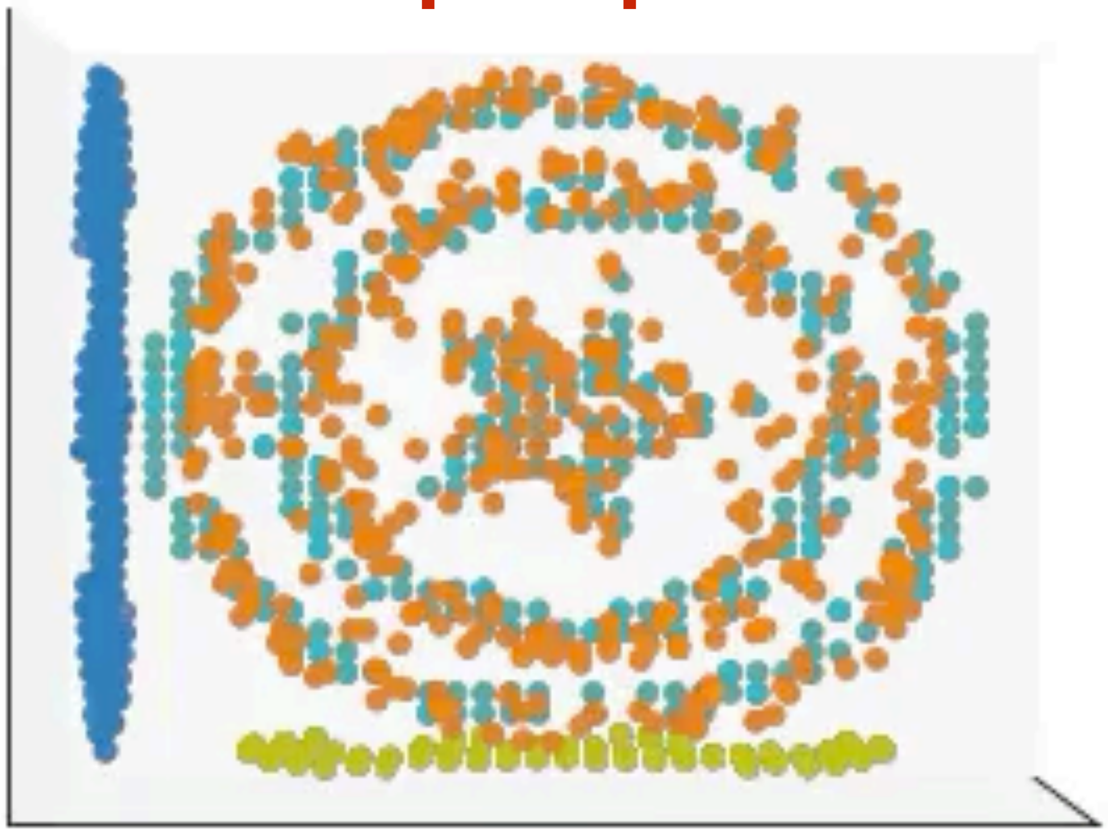
Prop [Tanguy et al. 2024]:

Barycenters are fixed points of G

+

Convergence $\rho_{n+1} = G(\rho_n)$

Works also with transport plans !



THANKS !