

Computing Optimal Transport Barycentres

Eloi Tanguy, Julie Delon, Nathaël Gozlan

MAP5, Université Paris-Cité

June 25, 2025



① Optimal Transport

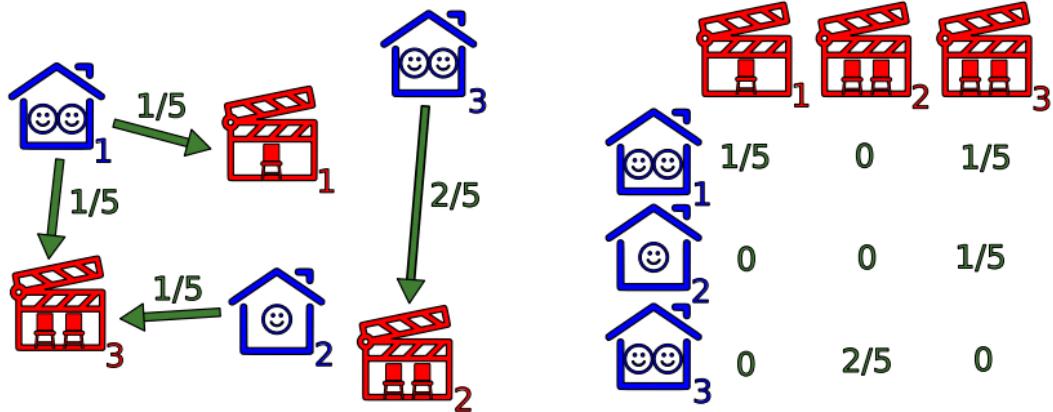
② Wasserstein Barycentres

③ OT Barycentres

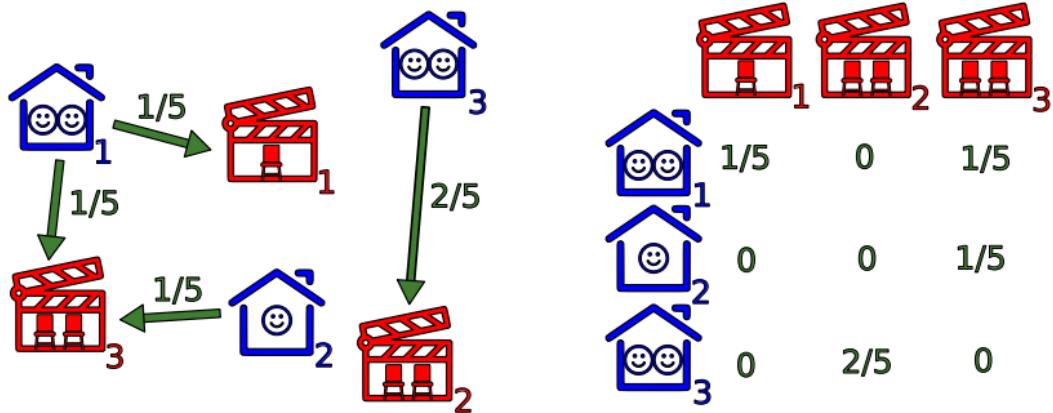
④ Discrete Case and Numerics

⑤ Application to GMMs

Discrete Optimal Transport



Discrete Optimal Transport



Assignment Cost:

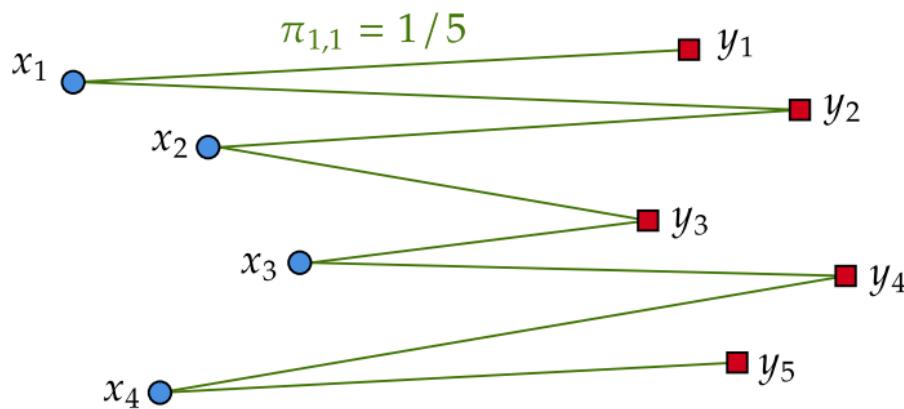
$$\frac{1}{5} \times c(x_1, y_1) + \frac{1}{5} \times c(x_1, y_3) + \frac{1}{5} \times c(x_2, y_3) + \frac{2}{5} \times c(x_3, y_2).$$

Constraints on $\pi \in \mathbb{R}_+^{3 \times 3}$: $\pi \mathbf{1} = (2/5, 1/5, 2/5)$, $\pi^\top \mathbf{1} = (1/5, 2/5, 2/5)$.Optimal Transport Cost : $\min_{\pi} \sum_{i,j} c(x_i, y_j) \pi_{i,j}$.

OT between discrete measures

$$\mu = \sum_{i=1}^n a_i \delta_{x_i}, \quad \nu = \sum_{j=1}^m b_j \delta_{y_j}$$

$$\mathcal{T}_c(\mu, \nu) = \inf_{\pi \in \Pi(a, b)} \sum_{i,j} c(x_i, y_j) \pi_{i,j}.$$



OT Cost and 2-Wasserstein Distance

$$\mathcal{T}_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) d\pi(x, y) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[c(X, Y)].$$

$$W_2^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} \|x - y\|_2^2 d\pi(x, y).$$

OT Cost and 2-Wasserstein Distance

$$\mathcal{T}_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) d\pi(x, y) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[c(X, Y)].$$

$$W_2^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} \|x - y\|_2^2 d\pi(x, y).$$

Bures-Wasserstein

$$\begin{aligned} W_2^2(\mathcal{N}(m_1, S_1), \mathcal{N}(m_2, S_2)) \\ = \|m_1 - m_2\|_2^2 \\ + \text{Tr} \left(S_1 + S_2 - 2(S_1^{1/2} S_2 S_1^{1/2})^{1/2} \right) \end{aligned}$$

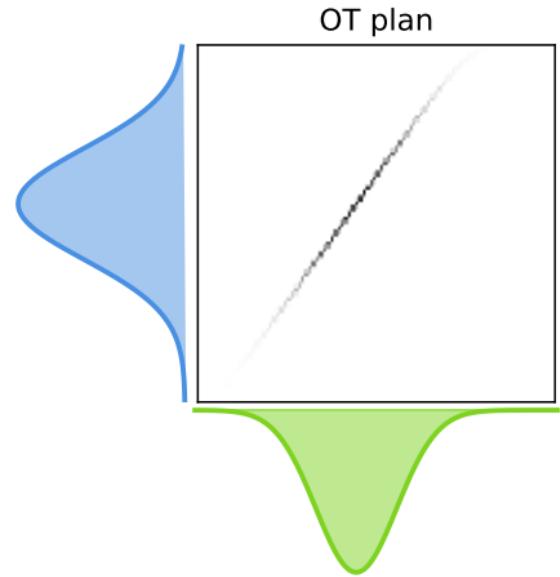
OT Cost and 2-Wasserstein Distance

$$\mathcal{T}_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) d\pi(x, y) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[c(X, Y)].$$

$$W_2^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} \|x - y\|_2^2 d\pi(x, y).$$

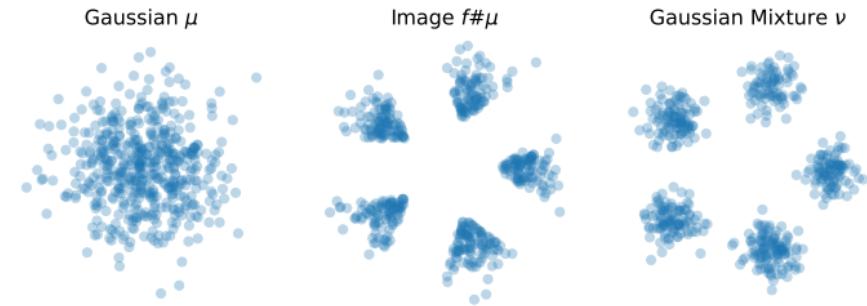
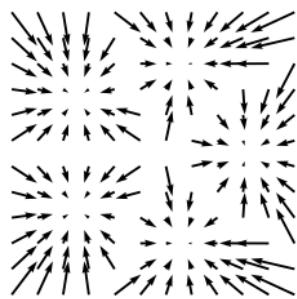
Bures-Wasserstein

$$\begin{aligned} W_2^2(\mathcal{N}(m_1, S_1), \mathcal{N}(m_2, S_2)) \\ = \|m_1 - m_2\|_2^2 \\ + \text{Tr} \left(S_1 + S_2 - 2(S_1^{1/2} S_2 S_1^{1/2})^{1/2} \right) \end{aligned}$$



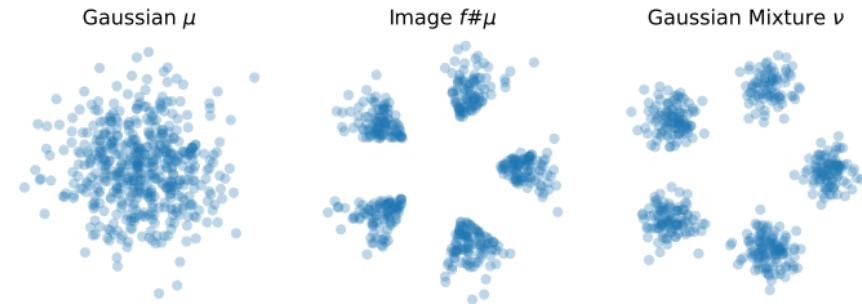
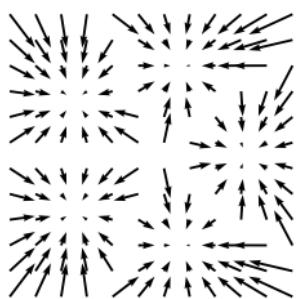
Push-forward measures and OT maps

Image Measure: $f\#\mu := \text{Law}_{X \sim \mu}[f(X)]$



Push-forward measures and OT maps

Image Measure: $f\#\mu := \text{Law}_{X \sim \mu}[f(X)]$

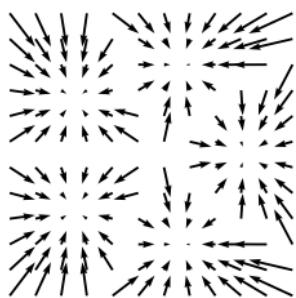
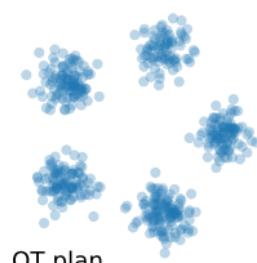
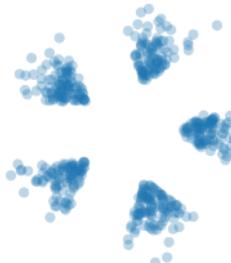
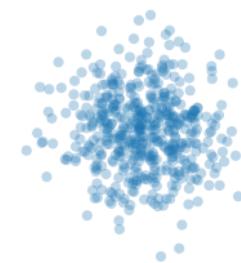


Brenier's Theorem

If $c(x, y) = \|x - y\|_2^2$, and $\mu \ll \mathcal{L}^d$, then there is a unique solution $\pi^* = (I, \nabla \varphi)\#\mu$, with φ convex.

Push-forward measures and OT maps

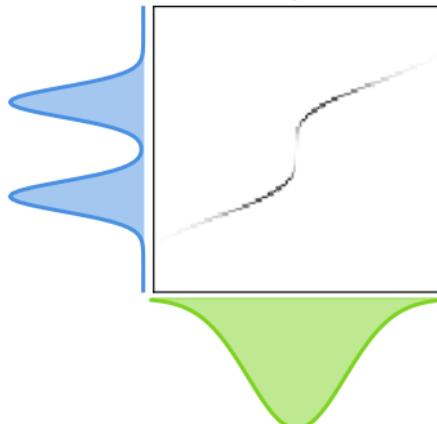
Image Measure: $f\#\mu := \text{Law}_{X \sim \mu}[f(X)]$

Gaussian μ Image $f\#\mu$ Gaussian Mixture ν 

OT plan

Brenier's Theorem

If $c(x, y) = \|x - y\|_2^2$, and $\mu \ll \mathcal{L}^d$, then there is a unique solution $\pi^* = (I, \nabla \varphi)\#\mu$, with φ convex.



① Optimal Transport

② Wasserstein Barycentres

③ OT Barycentres

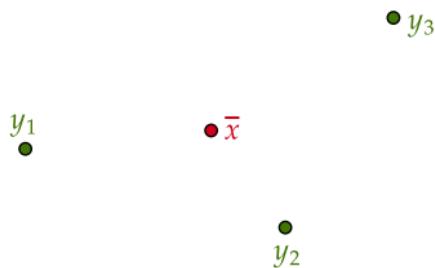
④ Discrete Case and Numerics

⑤ Application to GMMs

From Euclidean Combinations to Fréchet Means

$$\bar{x} = \sum_{k=1}^K \lambda_k y_k$$

$$\bar{x} = \operatorname{argmin}_{x \in \mathbb{R}^d} \sum_{k=1}^K \lambda_k \|x - y_k\|_2^2$$



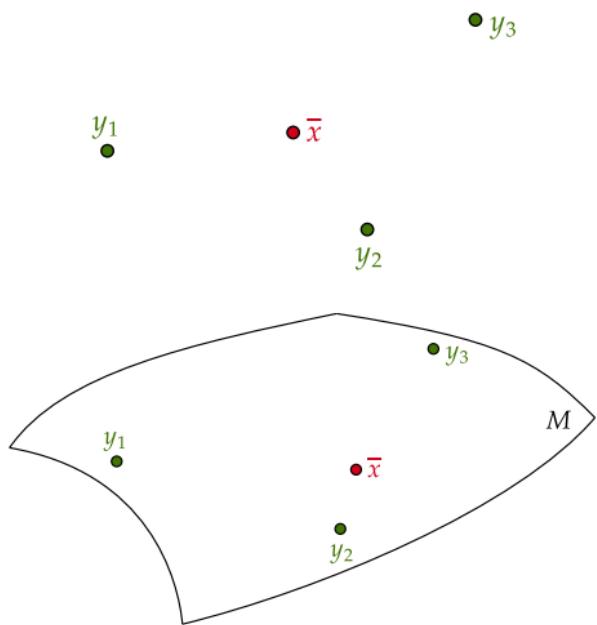
From Euclidean Combinations to Fréchet Means

$$\bar{x} = \sum_{k=1}^K \lambda_k y_k$$

$$\bar{x} = \operatorname{argmin}_{x \in \mathbb{R}^d} \sum_{k=1}^K \lambda_k \|x - y_k\|_2^2$$

Fréchet mean:

$$\bar{x} \in \operatorname{argmin}_{x \in \mathcal{X}} \sum_{k=1}^K \lambda_k d(x, y_k)^2.$$



From Euclidean Combinations to Fréchet Means

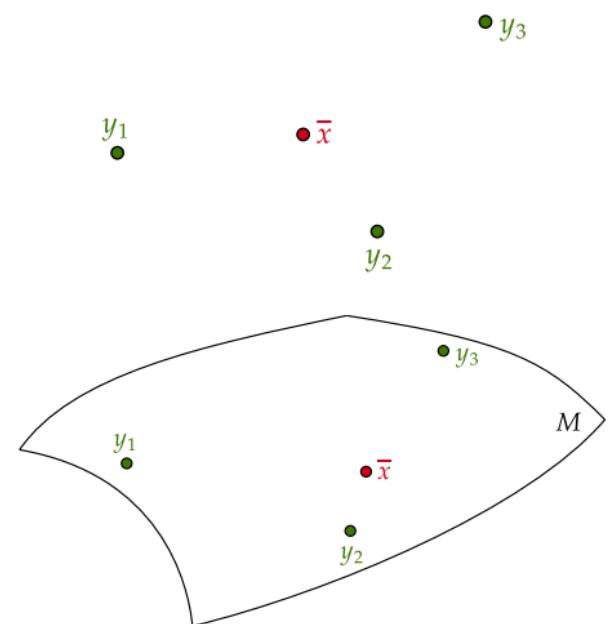
$$\bar{x} = \sum_{k=1}^K \lambda_k y_k$$

$$\bar{x} = \operatorname{argmin}_{x \in \mathbb{R}^d} \sum_{k=1}^K \lambda_k \|x - y_k\|_2^2$$

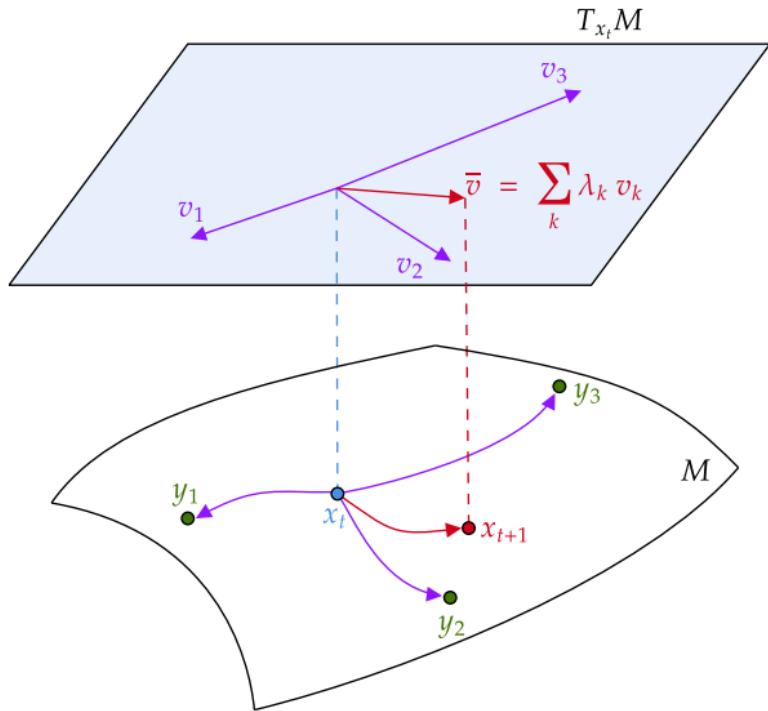
Fréchet mean:

$$\bar{x} \in \operatorname{argmin}_{x \in \mathcal{X}} \sum_{k=1}^K \lambda_k d(x, y_k)^2.$$

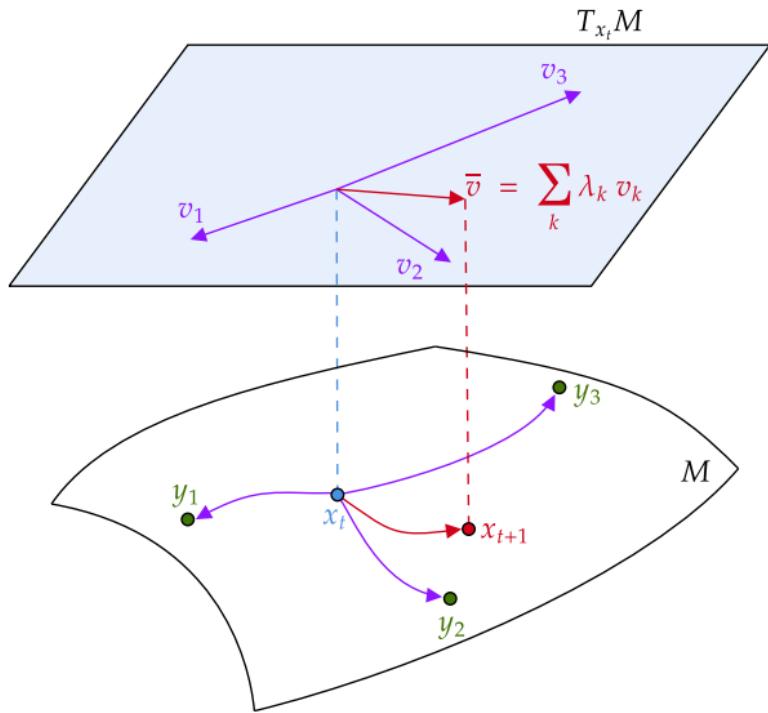
Generalisation: $\bar{x} \in \operatorname{argmin}_{x \in \mathcal{X}} \sum_{k=1}^K c_k(x, y_k).$



Fixed-Point Algorithm for Fréchet Means on Manifolds



Fixed-Point Algorithm for Fréchet Means on Manifolds



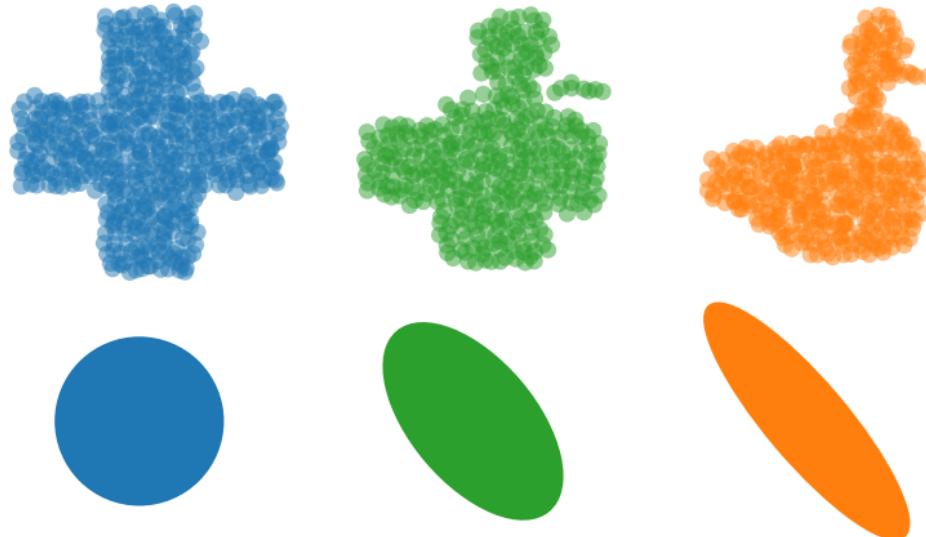
$$V(x) = \sum_{k=1}^K \lambda_k d(x, y_k)^2.$$

$$\nabla V(x) = -2 \sum_{k=1}^K \lambda_k \text{Log}_x(y_k).$$

$$x_{t+1} = \text{Exp}_{x_t} \left(-\frac{1}{2} \nabla V(x_t) \right).$$

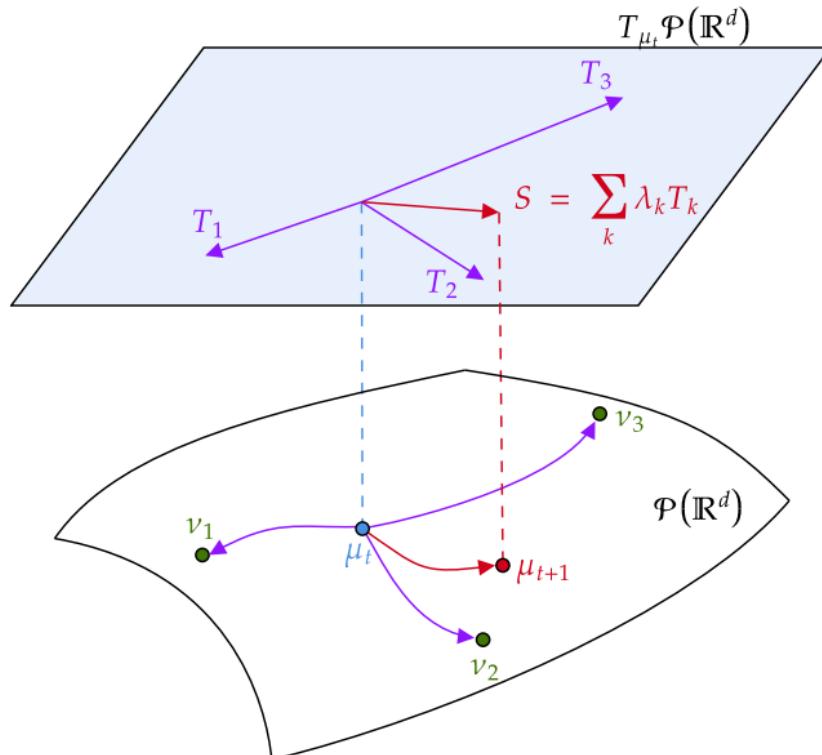
2-Wasserstein Barycentres (Aguech & Carlier 2011 [1])

$$\operatorname{argmin}_{\mu \in \mathcal{P}(\mathbb{R}^d)} \sum_{k=1}^K \lambda_k W_2^2(\mu, \nu_k).$$



Fixed-Point Method (Alvarez-Esteban et al. 2016 [3])

Assumptions: $c(x, y) = \|x - y\|_2^2$, AC measures on \mathbb{R}^d .



① Optimal Transport

② Wasserstein Barycentres

③ OT Barycentres

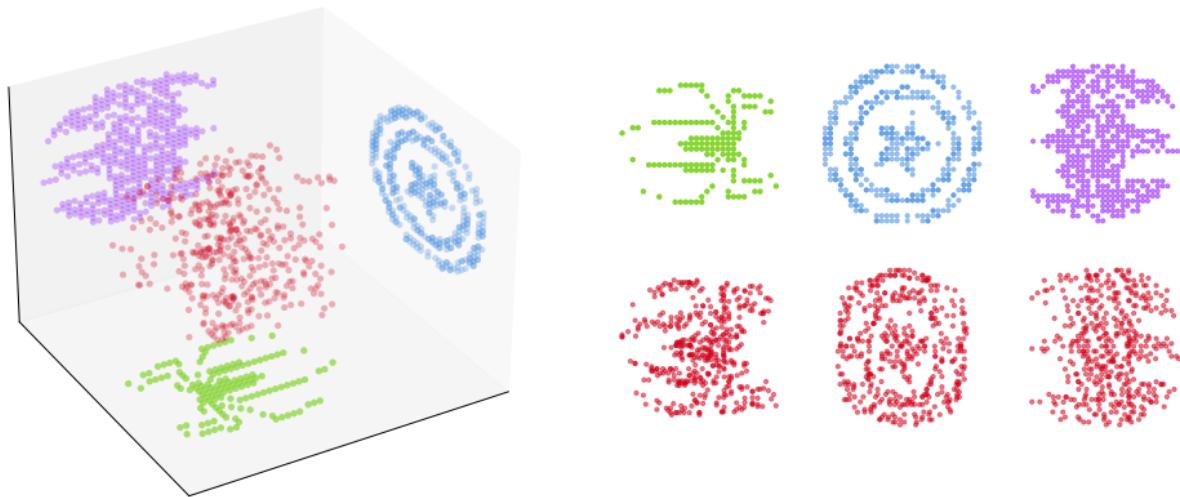
④ Discrete Case and Numerics

⑤ Application to GMMs

Motivation for OT barycenters with generic costs

$$W_1(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int \|x - y\|_2 d\pi(x, y).$$

Find $\mu \in \mathcal{P}(\mathbb{R}^3)$ minimising $\sum_k \frac{1}{3} W_1(P_k \# \mu, \nu_k)$ where $\nu_k \in \mathcal{P}(\mathbb{R}^2)$.



Generalises Delon et al. 2021 [5] where $c_k(x, y) = \|P_k(x) - y\|_2^2$.

Generalising Wasserstein Barycentres

Setting:

- $(\mathcal{X}, d_{\mathcal{X}})$ compact metric space for barycentres.
- $(\mathcal{Y}_k, d_{\mathcal{Y}_k})$ compact metric spaces for measures ν_k .
- $c_k : \mathcal{X} \times \mathcal{Y}_k \longrightarrow \mathbb{R}_+$ continuous cost functions.

Generalising Wasserstein Barycentres

Setting:

- $(\mathcal{X}, d_{\mathcal{X}})$ compact metric space for barycentres.
- $(\mathcal{Y}_k, d_{\mathcal{Y}_k})$ compact metric spaces for measures ν_k .
- $c_k : \mathcal{X} \times \mathcal{Y}_k \longrightarrow \mathbb{R}_+$ continuous cost functions.

$$\operatorname{argmin}_{\mu \in \mathcal{P}(\mathcal{X})} V(\mu), \quad V(\mu) := \sum_{k=1}^K \mathcal{T}_{c_k}(\mu, \nu_k).$$

Generalising Wasserstein Barycentres

Setting:

- $(\mathcal{X}, d_{\mathcal{X}})$ compact metric space for barycentres.
- $(\mathcal{Y}_k, d_{\mathcal{Y}_k})$ compact metric spaces for measures ν_k .
- $c_k : \mathcal{X} \times \mathcal{Y}_k \longrightarrow \mathbb{R}_+$ continuous cost functions.

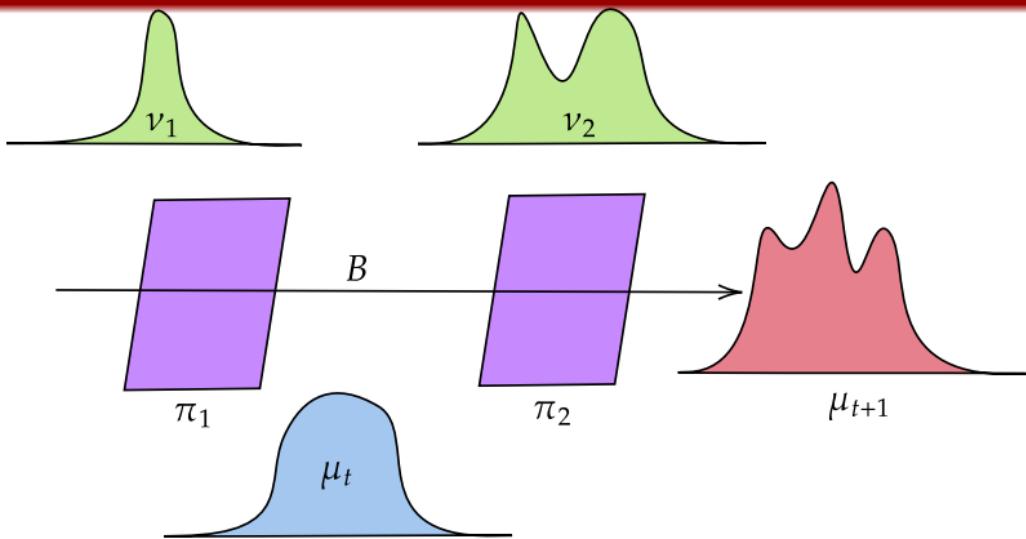
$$\underset{\mu \in \mathcal{P}(\mathcal{X})}{\operatorname{argmin}} V(\mu), \quad V(\mu) := \sum_{k=1}^K \mathcal{T}_{c_k}(\mu, \nu_k).$$

Assumption: The ground barycenter function

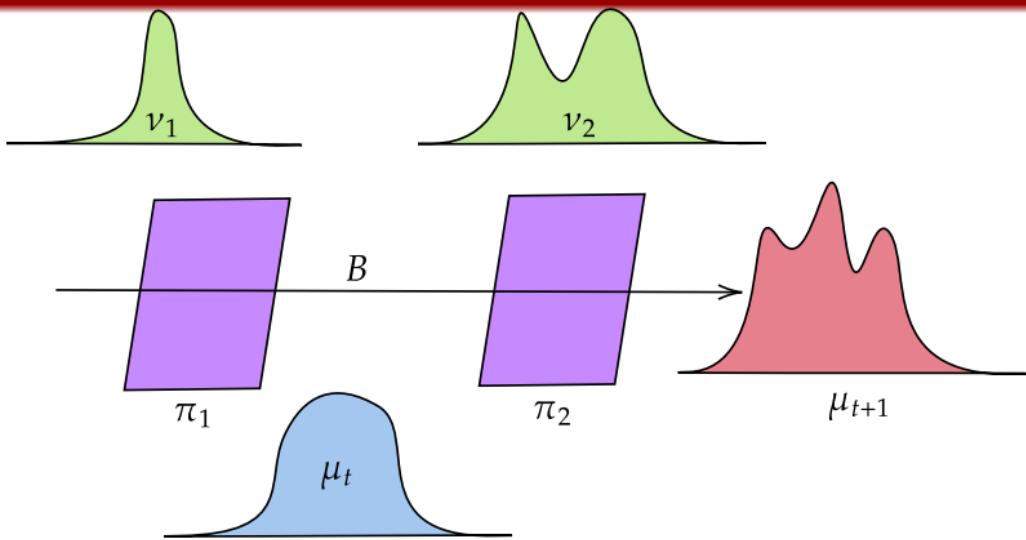
$$B(y_1, \dots, y_K) := \underset{x \in \mathcal{X}}{\operatorname{argmin}} \sum_{k=1}^K c_k(x, y_k)$$

is well-defined.

Fixed-Point Algorithm: Intuition



Fixed-Point Algorithm: Intuition

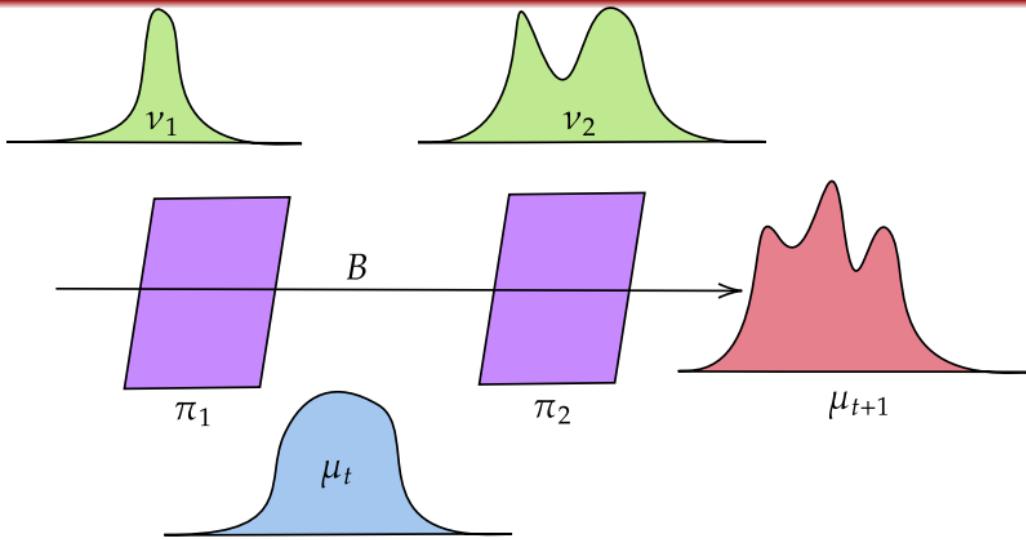


General Idea

Let (X_t, Y_1, \dots, Y_K) RVs such that $X_t \sim \mu_t$, $Y_k \sim \nu_k$ and $(X_t, Y_k) \sim \pi_k \in \Pi_{c_k}^*(\mu_t, \nu_k)$. Take $X_{t+1} = B(Y_1, \dots, Y_K)$.

If $\Pi_{c_k}^*(\mu_t, \nu_k) = \{(I, T_k)\# \mu_t\}$ then $\mu_{t+1} = B(T_1, \dots, T_K)\# \mu_t$.

Fixed-point Algorithm: (more) formal definition



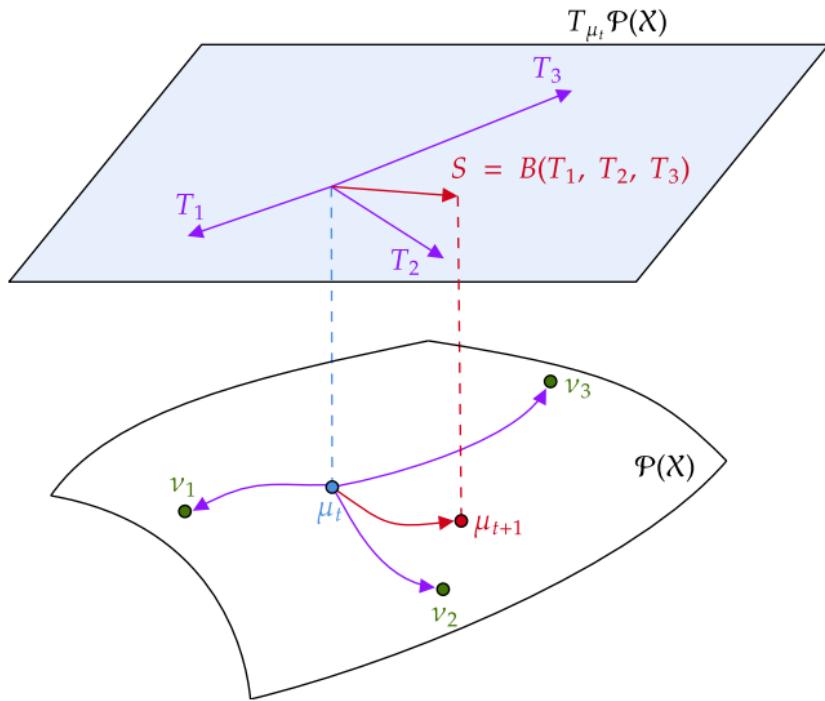
$$\Gamma(\mu) := \left\{ \gamma \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}_1 \cdots \times \mathcal{Y}_K) : \forall k \in \llbracket 1, K \rrbracket, \gamma_{0,k} \in \Pi_{c_k}^*(\mu, \nu_k) \right\},$$

$$G := \begin{cases} \mathcal{P}(\mathcal{X}) & \rightrightarrows \mathcal{P}(\mathcal{X}) \\ \mu & \mapsto B \# \Gamma(\mu) \end{cases}, \quad \mu_{t+1} \in G(\mu_t).$$

$$B \# \Gamma(\mu) := \{B \# \gamma, \gamma \in \Gamma(\mu)\}, \quad B \# \gamma = \text{Law}_{(X, Y_1, \dots, Y_K) \sim \gamma} B(Y_1, \dots, Y_K).$$

Relation to Alvarez-Esteban et al. 2016 [3]

Dream case: $\mathcal{X} = \mathcal{Y}_1 = \dots = \mathcal{Y}_K$ and maps exist.



$$\gamma : \gamma_{0,k} \in \Pi_{c_k}^*(\mu_t, \nu_k),$$

$$\mu_{t+1} = B\#\gamma.$$

Algorithm Convergence

Ground Barycentre Lemma

$$\sum_k c_k(x, y_k) \geq \sum_k c_k(B(y_1, \dots, y_K), y_k) + \delta(x, B(y_1, \dots, y_K)).$$

Case $\|x - y\|_2^2$: simply $\sum_k \lambda_k \|x - y_k\|_2^2 = \sum_k \|\bar{x} - y_k\|_2^2 + \|x - \bar{x}\|_2^2$.

Algorithm Convergence

Ground Barycentre Lemma

$$\sum_k c_k(x, y_k) \geq \sum_k c_k(B(y_1, \dots, y_K), y_k) + \delta(x, B(y_1, \dots, y_K)).$$

Case $\|x - y\|_2^2$: simply $\sum_k \lambda_k \|x - y_k\|_2^2 = \sum_k \|\bar{x} - y_k\|_2^2 + \|x - \bar{x}\|_2^2$.

Decrease Property

$$\forall \bar{\mu} \in G(\mu), V(\mu) \geq V(\bar{\mu}) + \mathcal{T}_\delta(\mu, \bar{\mu}).$$

If μ^* is a barycentre then $G(\mu^*) = \{\mu^*\}$.

Algorithm Convergence

Ground Barycentre Lemma

$$\sum_k c_k(x, y_k) \geq \sum_k c_k(B(y_1, \dots, y_K), y_k) + \delta(x, B(y_1, \dots, y_K)).$$

Case $\|x - y\|_2^2$: simply $\sum_k \lambda_k \|x - y_k\|_2^2 = \sum_k \|\bar{x} - y_k\|_2^2 + \|x - \bar{x}\|_2^2$.

Decrease Property

$$\forall \bar{\mu} \in G(\mu), V(\mu) \geq V(\bar{\mu}) + \mathcal{T}_\delta(\mu, \bar{\mu}).$$

If μ^* is a barycentre then $G(\mu^*) = \{\mu^*\}$.

Convergence

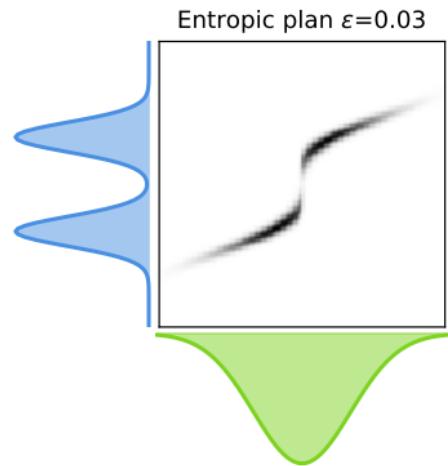
If μ is a subsequential limit of (μ_t) then $\mu \in G(\mu)$.

Entropic Barycentres

$$\mathcal{T}_{c,\varepsilon}(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} cd\pi + \varepsilon \text{KL}(\pi | \mu \otimes \nu).$$

$$V_\varepsilon(\mu) := \sum_{k=1}^K \mathcal{T}_{c,\varepsilon}(\mu, \nu_k).$$

$$G_\varepsilon(\mu) := B \# \gamma, \text{ with } \gamma_{0,k} = \Pi_{c_k, \varepsilon}^*(\mu, \nu_k).$$

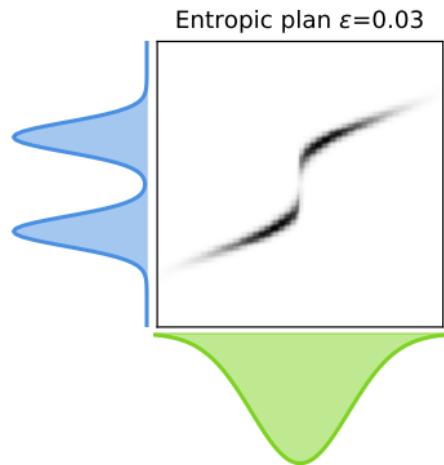


Entropic Barycentres

$$\mathcal{T}_{c,\varepsilon}(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} cd\pi + \varepsilon \text{KL}(\pi | \mu \otimes \nu).$$

$$V_\varepsilon(\mu) := \sum_{k=1}^K \mathcal{T}_{c,\varepsilon}(\mu, \nu_k).$$

$$G_\varepsilon(\mu) := B \# \gamma, \text{ with } \gamma_{0,k} = \Pi_{c_k, \varepsilon}^*(\mu, \nu_k).$$



Decrease Property

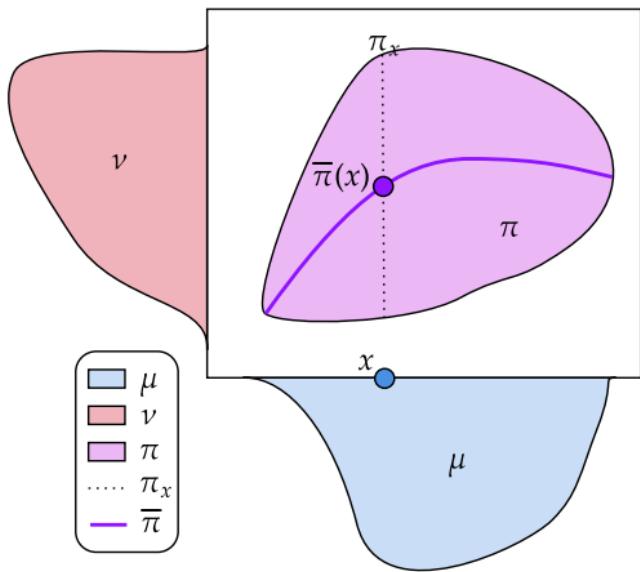
$V_\varepsilon(\mu) \geq V_\varepsilon(G_\varepsilon(\mu)) + \mathcal{T}_\delta(\mu, G_\varepsilon(\mu)).$ If μ^* barycentre, $G_\varepsilon(\mu^*) = \mu^*.$

Convergence

If μ is a subsequential limit of (μ_t) then $\mu = G_\varepsilon(\mu).$

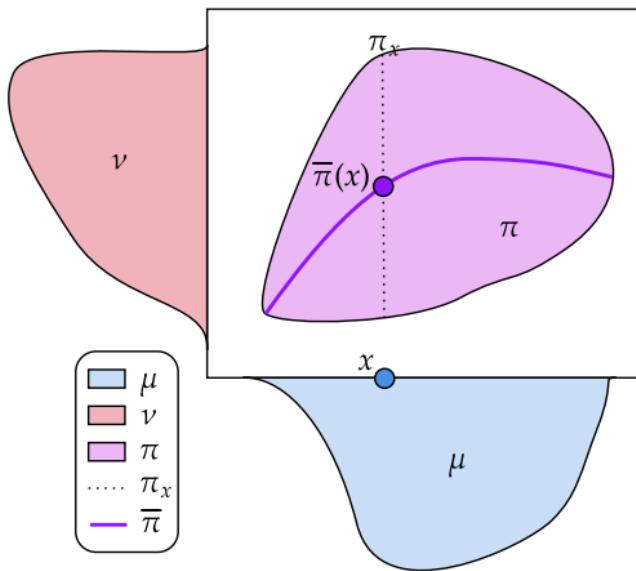
Barycentric Projections

Replace a coupling π with a map $\bar{\pi}$.



Barycentric Projections

Replace a coupling π with a map $\bar{\pi}$.



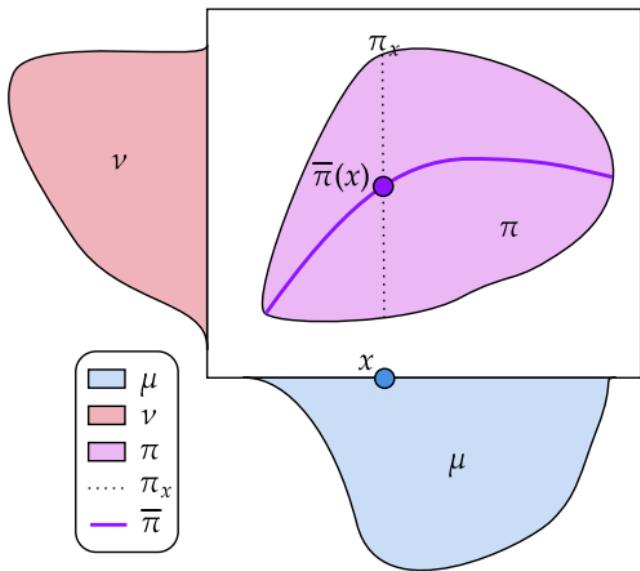
$$\bar{\pi}(x) = \int y d\pi_x(y).$$

$$\bar{\pi}(x) = \mathbb{E}_{(X,Y) \sim \pi}[Y | X = x].$$

$$\bar{\pi} = \operatorname{argmin}_{f \in L^2(\mu)} \int \|f(x) - y\|_2^2 d\pi(x, y).$$

Barycentric Projections

Replace a coupling π with a map $\bar{\pi}$.



$$\bar{\pi}(x) = \int y d\pi_x(y).$$

$$\bar{\pi}(x) = \mathbb{E}_{(X,Y) \sim \pi}[Y | X = x].$$

$$\bar{\pi} = \operatorname{argmin}_{f \in L^2(\mu)} \int \|f(x) - y\|_2^2 d\pi(x, y).$$

$$H(\mu) = \left\{ B(\bar{\pi}_1, \dots, \bar{\pi}_K) \# \mu, \pi_k \in \Pi_{c_k}^*(\mu, \nu_k) \right\}.$$



No guarantees.

① Optimal Transport

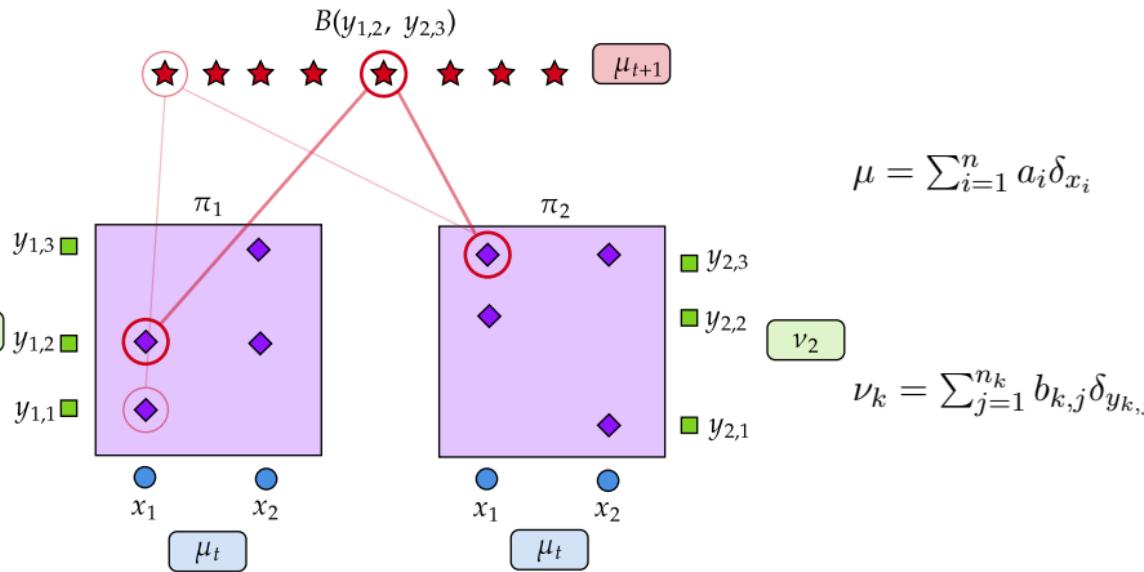
② Wasserstein Barycentres

③ OT Barycentres

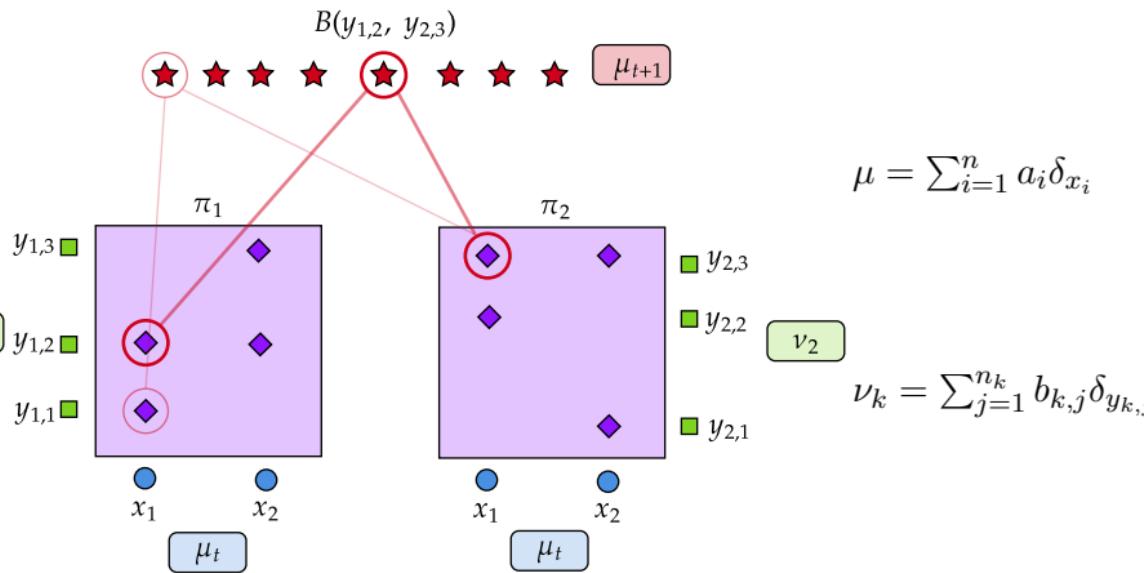
④ Discrete Case and Numerics

⑤ Application to GMMs

Discrete G



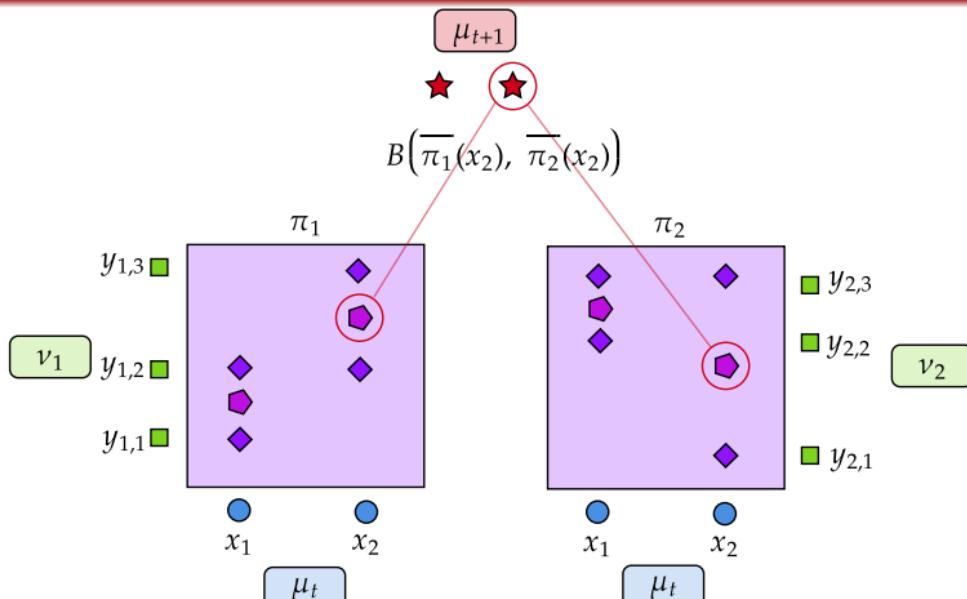
Discrete G



$$G(\mu) = \left\{ \sum_{j_1, \dots, j_K} \left(\sum_{i=1}^n \frac{1}{a_i^{K-1}} \pi_{i,j_1}^{(1)} \times \dots \times \pi_{i,j_K}^{(K)} \right) \delta(B(y_{1,j_1}, \dots, y_{K,j_K})) , \right.$$

$$\left. \pi^{(k)} \in \Pi_{c_k}^*(\mu, \nu_k) \right\}.$$

Discrete H (Generalises Cuturi & Doucet 2014 [4])



$$H(\mu) = \left\{ \sum_{i=1}^n a_i \delta(B(\bar{\pi}_1(x_i), \dots, \bar{\pi}_K(x_i))), \pi_k \in \Pi_{c_k}^*(\mu, \nu_k) \right\},$$

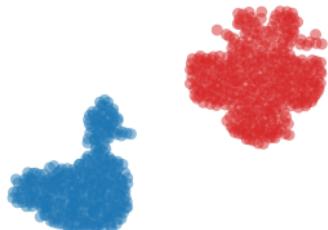
$$\bar{\pi}_k(x_i) = (1/a_i) \sum_{j=1}^{n_1} \pi_{i,j}^{(k)} y_{1,j}.$$

Illustration for $c(x, y) = \|x - y\|_{1.5}^{1.5}$

Barycentre for the cost $|x - y|_{3/2}^{3/2}$



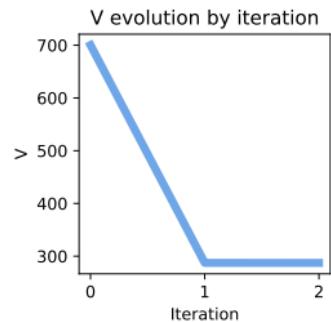
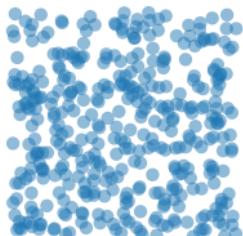
Iteration 0



Iteration 1



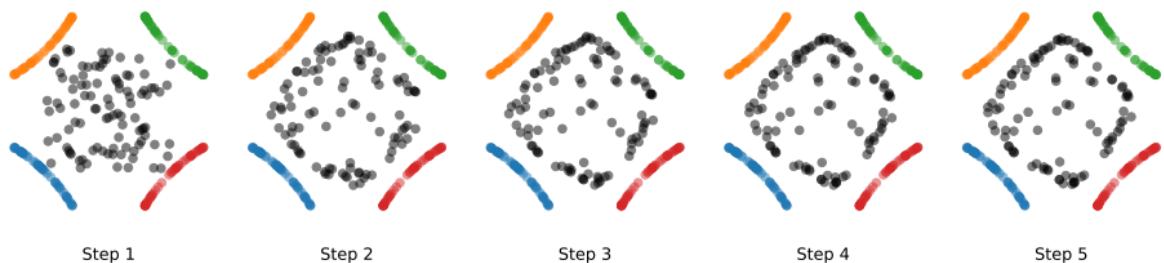
Iteration 2



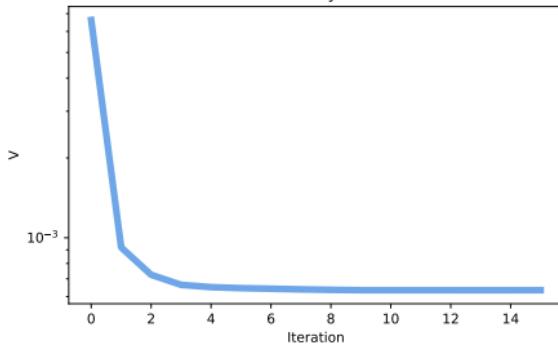
Non-linear Generalised Wasserstein Barycentre

$\operatorname{argmin}_{\mu} \sum_{k=1}^4 \frac{1}{4} W_2^2(P_k \# \mu, \nu_k)$ where P_k is the projection onto circle k .

First 5 Steps Fixed-point GWB solver



V evolution by iteration



① Optimal Transport

② Wasserstein Barycentres

③ OT Barycentres

④ Discrete Case and Numerics

⑤ Application to GMMs

OT between GMMs

$$W_2^2(\mathcal{N}(m_1, S_1), \mathcal{N}(m_2, S_2)) = \|m_1 - m_2\|_2^2 + \underbrace{\text{Tr} \left(S_1 + S_2 - 2(S_1^{1/2} S_2 S_1^{1/2})^{1/2} \right)}_{d_{\text{BW}}^2(S_1, S_2) :=}$$

OT between GMMs

$$W_2^2(\mathcal{N}(m_1, S_1), \mathcal{N}(m_2, S_2)) = \|m_1 - m_2\|_2^2 + \underbrace{\text{Tr} \left(S_1 + S_2 - 2(S_1^{1/2} S_2 S_1^{1/2})^{1/2} \right)}_{d_{\text{BW}}^2(S_1, S_2) :=}$$

Ground space: $(\mathcal{X}, d) = (\mathcal{Y}_k, d_{\mathcal{Y}_k}) = (\mathcal{N}, W_2)$ with ground cost $c = W_2^2$.

$$\mu = \sum_{i=1}^n a_i \delta_{\mathcal{N}(m_i, S_i)}, \quad \nu = \sum_{j=1}^m b_j \delta_{\mathcal{N}(m'_j, S'_j)} \in \mathcal{P}(\mathcal{N});$$

$$\mathcal{T}_{W_2^2}(\mu, \nu) = \min_{\pi \in \Pi(a, b)} \sum_{i,j} (\|m_i - m'_j\|_2^2 + d_{\text{BW}}^2(S_i, S'_j)) \pi_{i,j}.$$

Ground Barycentre Between Gaussians

Gaussian barycentres (Aguech & Carlier 2011 [1]).

$$B(\mathcal{N}(m_1, S_1), \dots, \mathcal{N}(m_K, S_K)) = \mathcal{N}(\bar{m}, \bar{S}),$$

$$\bar{m} := \sum_{k=1}^K \lambda_k m_k, \quad \bar{S} := \operatorname{argmin}_{S \in S_d^{++}(\mathbb{R})} \sum_{k=1}^K \lambda_k d_{\text{BW}}^2(S, S_k).$$

Fixed-point computation for \bar{S} :

$$G_{\mathcal{N}}(S) = S^{-1/2} \left(\sum_{k=1}^K \lambda_k (S^{1/2} S_k S^{1/2})^{1/2} \right)^2 S^{-1/2}.$$

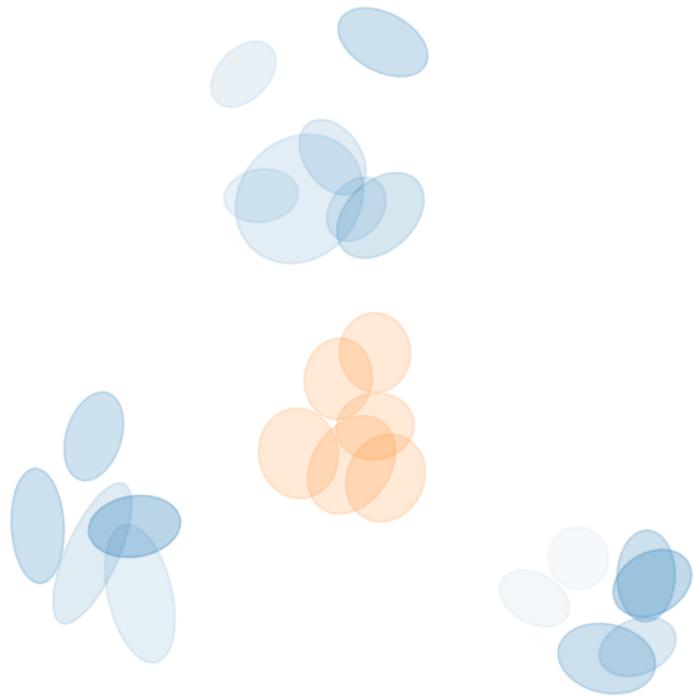
Riemannian gradient descent interpretation by Altschuler et al. 2021 [2].

GMM Barycentre

$$\mu = \sum_{i=1}^n a_i \delta_{\mathcal{N}(m_i, S_i)},$$

$$\nu_k = \sum_{j=1}^{n_k} b_k \delta_{\mathcal{N}(m_{k,j}, S_{k,j})},$$

$$V(\mu) = \sum_{k=1}^K \lambda_k \mathcal{T}_{W_2^2}(\mu, \nu_k).$$



GMM Barycentre Example



- Talk based on *ET, Julie Delon and Nathaël Gozlan (2024): Computing Barycentres of Measures for Generic Transport Costs.* arXiv preprint 2501.04016.
- All code at https://github.com/eloitanguy/ot_bar
- Functions (soon) released on <https://pythonot.github.io/>
- Slides at <https://eloitanguy.github.io/publications/>

Thanks!

- [1] Martial Aguech and Guillaume Carlier.
Barycenters in the Wasserstein space.
SIAM Journal on Mathematical Analysis, 43(2):904–924, 2011.
- [2] Jason Altschuler, Sinho Chewi, Patrik R Gerber, and Austin Stromme.
Averaging on the bures-wasserstein manifold: dimension-free convergence of gradient descent.
Advances in Neural Information Processing Systems, 34:22132–22145, 2021.
- [3] Pedro C Álvarez-Esteban, E Del Barrio, JA Cuesta-Albertos, and C Matrán.
A fixed-point approach to barycenters in Wasserstein space.
Journal of Mathematical Analysis and Applications, 441(2):744–762, 2016.

[4] Marco Cuturi and Arnaud Doucet.

Fast computation of Wasserstein barycenters.

In Eric P. Xing and Tony Jebara, editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 685–693, Bejing, China, 22–24 Jun 2014. PMLR.

[5] Julie Delon, Nathaël Gozlan, and Alexandre Saint-Dizier.

Generalized Wasserstein barycenters between probability measures living on different subspaces, 2021.