

Optimal transport

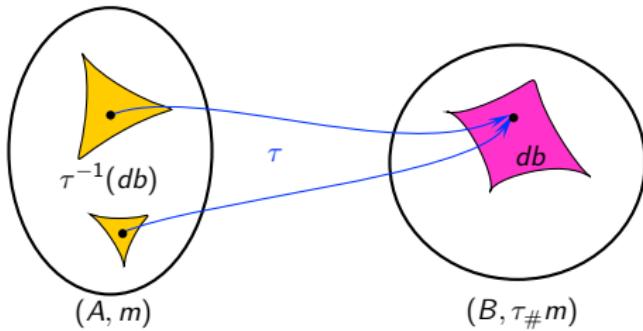
Christian Léonard

Université Paris Ouest

Séminaire MME-DII - Cergy - 2013, June

Push-forward of a measure

- $\mathcal{P}(A) = \{\text{probability measures on } A\}$
- $\tau : A \rightarrow B$ measurable mapping

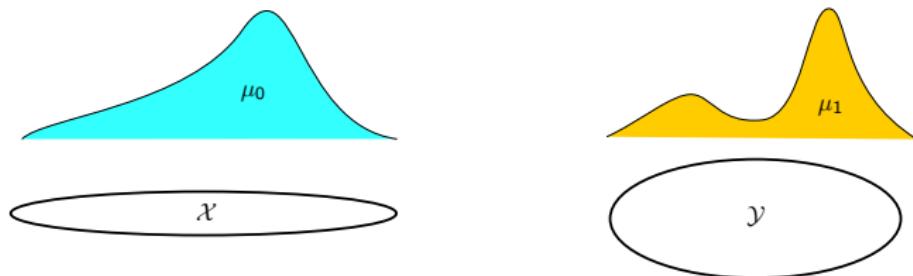


Push-forward of $m \in \mathcal{P}(A)$ by the mapping τ

- $m_{\#}\tau(db) := m(\tau^{-1}(db))$
- $m_{\#}\tau \in \mathcal{P}(B)$

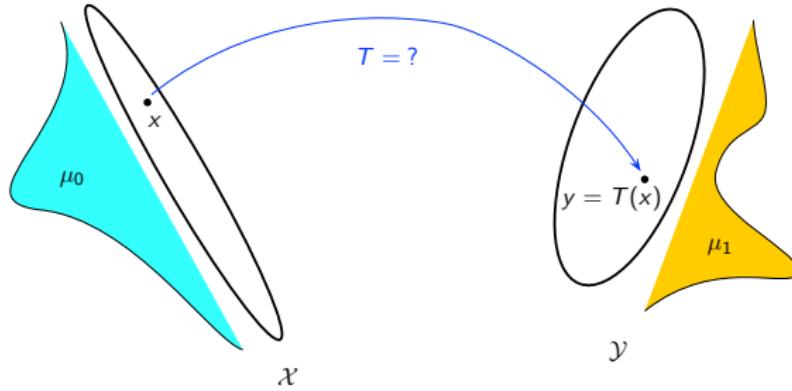
Monge problem

- \mathcal{X}, \mathcal{Y} state spaces
 - $\mathcal{X}, \mathcal{Y} = \mathbb{R}^n$, manifold, discrete space
- $c : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty)$ cost function
 - $c(x, y)$ transport cost of a unit mass from $x \in \mathcal{X}$ to $y \in \mathcal{Y}$
 - $\mathcal{X} = \{x_1, \dots, x_r\}, \mathcal{Y} = \{y_1, \dots, y_c\}, c = (c_{ij})_{1 \leq i \leq r, 1 \leq j \leq c}$
 - $\mathcal{X} = \mathcal{Y}$ metric space (\mathcal{X}, d)
 - ★ $c = d$
 - ★ $c = d^2$
 - $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n, c(x, y) = c(|y - x|)$ with c convex or concave
 - $\mu_0 \in \mathcal{P}(\mathcal{X}), \mu_1 \in \mathcal{P}(\mathcal{Y})$ prescribed probability measures



Monge problem

- $c : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty)$, $\mu_0 \in \mathcal{P}(\mathcal{X})$, $\mu_1 \in \mathcal{P}(\mathcal{Y})$ are given
- Find an optimal transport map $T : \mathcal{X} \rightarrow \mathcal{Y}$



Monge problem

$$\int_{\mathcal{X}} c(x, T(x)) \mu_0(dx) \rightarrow \min; \quad T : \mathcal{X} \rightarrow \mathcal{Y} : T_{\#}\mu_0 = \mu_1$$

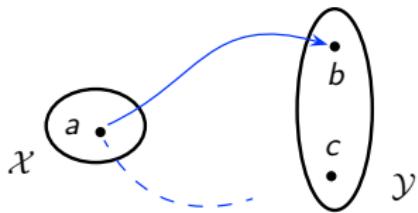
Monge problem

Monge problem

$$\int_{\mathcal{X}} c(x, T(x)) \mu_0(dx) \rightarrow \min; \quad T : \mathcal{X} \rightarrow \mathcal{Y} : T_{\#}\mu_0 = \mu_1$$

It is a difficult *nonlinear* problem

- $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n, \quad \mu_0(dx) = \mu_0(x) dx, \quad \mu_1(dy) = \mu_1(y) dy$
If T is differentiable,
 $T_{\#}\mu_0 = \mu_1 \iff \mu_0(x) = \mu_1(T(x)) |\det \nabla T(x)|, \forall x$
- No solution in general
 $\mathcal{X} = \{a\}, \quad \mathcal{Y} = \{b, c\}, \quad \mu_0 = \delta_a, \quad \mu_1 = (\delta_b + \delta_c)/2$



Coupling

Coupling of (μ_0, μ_1)

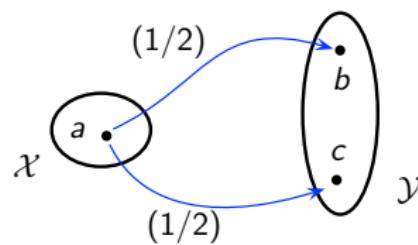
- ① $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ such that

$$\begin{cases} \pi_0(dx) := (X_0)_\# \pi(dx) = \pi(dx \times \mathcal{Y}) = \mu_0(dx) \\ \pi_1(dy) := (X_1)_\# \pi(dy) = \pi(\mathcal{X} \times dy) = \mu_1(dy) \end{cases}$$

- ② π is also called a *transport plan* between μ_0 and μ_1

- Stupid coupling: $\pi = \mu_0 \otimes \mu_1$
- Deterministic coupling: $\pi^T(dxdy) = \mu_0(dx)\delta_{T(x)}(dy)$, $\mu_1 = T_\#\mu_0$
The support of π^T is included in the *graph* of T
- $\mathcal{X} = \{a\}$, $\mathcal{Y} = \{b, c\}$, $\mu_0 = \delta_a$, $\mu_1 = (\delta_b + \delta_c)/2$

$$\pi = \delta_a \otimes (\delta_b + \delta_c)/2$$



Monge-Kantorovich problem

Monge problem

$$\int_{\mathcal{X}} c(x, T(x)) \mu_0(dx) \rightarrow \min; \quad T : \mathcal{X} \rightarrow \mathcal{Y} : T_{\#}\mu_0 = \mu_1$$

Monge-Kantorovich problem

$$\iint_{\mathcal{X} \times \mathcal{Y}} c(x, y) \pi(dx dy) \rightarrow \min; \quad \pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \pi_0 = \mu_0, \pi_1 = \mu_1$$

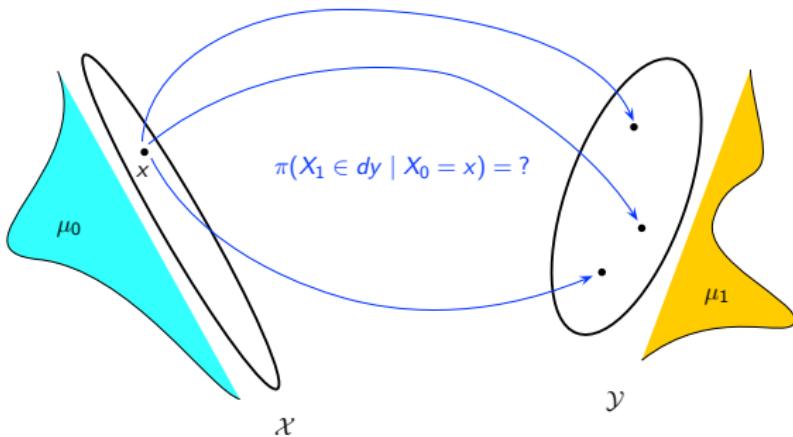
- $\pi \mapsto \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) \pi(dx dy) = \langle c, \pi \rangle$ is affine
- $\{\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \pi_0 = \mu_0, \pi_1 = \mu_1\}$ is a convex set
- Monge transport plan: $\pi^T(dx dy) = \mu_0(dx) \delta_{T(x)}(dy)$

- A *convex relaxation* of the nonlinear Monge problem
Infinite dimensional *linear programming* problem



Monge-Kantorovich problem

- $c : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty)$, $\mu_0 \in \mathcal{P}(\mathcal{X})$, $\mu_1 \in \mathcal{P}(\mathcal{Y})$ are given
- Find an optimal transport kernel $\pi(X_1 \in \cdot | X_0 = x) : \mathcal{X} \rightarrow \mathcal{P}(\mathcal{Y})$
Recall: $\pi(dxdy) = \mu_0(dx) \pi(X_1 \in dy | X_0 = x)$



Monge-Kantorovich problem

$$\int_{\mathcal{X} \times \mathcal{Y}} c(x, y) \pi(dxdy) \rightarrow \min; \quad \pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \pi_0 = \mu_0, \pi_1 = \mu_1 \quad (\text{MK})$$

Monge-Kantorovich problem

- Regular framework (don't mind if you're lost)
 - ▶ \mathcal{X}, \mathcal{Y} are Polish (complete metrizable, separable) spaces
 - ▶ probability measures are Borel measures
 - ▶ $c : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty)$ is lower semicontinuous
- Example: $\mathcal{X} = \mathcal{Y}$, $c(x, y) = \mathbf{1}_{\{x \neq y\}}$
- Existence

Typical existence result

Assume: $c(x, y) \leq c_0(x) + c_1(y)$, $\int_{\mathcal{X}} c_0 d\mu_0, \int_{\mathcal{Y}} c_1 d\mu_1 < \infty$.

Then, (MK) admits a solution (optimal plan).

Sketch of proof.

- ▶ $\int_{\mathcal{X} \times \mathcal{Y}} c d(\mu_0 \otimes \mu_1) \leq \int_{\mathcal{X}} c_0 d\mu_0 + \int_{\mathcal{Y}} c_1 d\mu_1 < \infty$
- ▶ tightness of $\{\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \pi_0 = \mu_0, \pi_1 = \mu_1\}$
- ▶ lower semicontinuity of $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \mapsto \int_{\mathcal{X} \times \mathcal{Y}} c d\pi$

□

- Uniqueness

Non-uniqueness is the rule

Wasserstein distance

- $\mathcal{X} = \mathcal{Y}$ metric space (\mathcal{X}, d)
- $c(x, y) = d^p(x, y), \quad 1 \leq p < \infty$
- $\mathcal{P}_p(\mathcal{X}) := \left\{ \mu \in \mathcal{P}(\mathcal{X}) : \int_{\mathcal{X}} d^p(x_o, x) \mu(dx) < \infty \right\}$

Theorem

For all $\mu_0, \mu_1 \in \mathcal{P}_p(\mathcal{X})$,

- ① (MK) admits a solution
- ② $W_p(\mu_0, \mu_1) := \inf (\text{MK})^{1/p} < \infty$ is a distance on $\mathcal{P}_p(\mathcal{X})$
(Wasserstein distance)

Sketch of proof.

- ▶ $c(x, y) = d^p(x, y) \leq [d(x, x_o) + d(x_o, y)]^p \leq 2^p[d^p(x_o, x) + d^p(x_o, y)]$
- ▶ previous theorem yields result 1)
- ▶ 2) triangle inequality for d , Hölder's inequality and Markov:
 $\pi_{01}^* = \text{Law}(X_0, X_1)$, $\pi_{12}^* = \text{Law}(X_1, X_2)$, $\pi_{02} = \text{Law}(X_0, X_2)$
gives $W_p(\mu_0, \mu_2) \leq W_p(\mu_0, \mu_1) + W_p(\mu_1, \mu_2)$ \square

Duality

Consider functions $\begin{cases} \varphi : \mathcal{X} \rightarrow [-\infty, \infty) \\ \psi : \mathcal{Y} \rightarrow [-\infty, \infty) \end{cases}$, $\varphi(x) + \psi(y) =: \varphi \oplus \psi(x, y)$

The dual problem of (MK)

$$\int_{\mathcal{X}} \varphi(x) \mu_0(dx) + \int_{\mathcal{Y}} \psi(y) \mu_1(dy) \rightarrow \max; \quad \varphi, \psi : \varphi \oplus \psi \leq c \quad (\text{D})$$

Characterization of optimal plan (informal)

Let $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ be such that $\pi_0 = \mu_0$, $\pi_1 = \mu_1$. TFAE

- ① π solves (MK)
- ② there exists (φ, ψ) such that $\begin{cases} \varphi \oplus \psi \leq c, \text{ everywhere} \\ \varphi \oplus \psi = c, \text{ } \pi\text{-everywhere} \end{cases}$

In such case, (φ, ψ) solves (D).

These (D)-optimal φ, ψ are called Kantorovich potentials

Duality

- Transport $\mu_0 \in \mathcal{P}(\mathcal{X})$ onto $\mu_1 \in \mathcal{P}(\mathcal{Y})$ at minimal cost \inf (MK)

$$\int_{\mathcal{X} \times \mathcal{Y}} c(x, y) \pi(dx dy) \rightarrow \min; \quad \pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \pi_0 = \mu_0, \pi_1 = \mu_1 \quad (\text{MK})$$

- Prices of a transport company:

- ▶ $\varphi(x)$ to take away a unit mass from x
- ▶ $\psi(y)$ to deliver a unit mass at y
- ▶ competitive prices since $\varphi(x) + \psi(y) \leq c(x, y)$

Duality

Dual problems

$$\int_{\mathcal{X} \times \mathcal{Y}} c(x, y) \pi(dx dy) \rightarrow \min; \quad \pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \pi_0 = \mu_0, \pi_1 = \mu_1 \quad (\text{MK})$$

$$\int_{\mathcal{X}} \varphi(x) \mu_0(dx) + \int_{\mathcal{Y}} \psi(y) \mu_1(dy) \rightarrow \max; \quad \varphi, \psi : \varphi \oplus \psi \leq c \quad (\text{D})$$

- (D) is the income maximization of the transport company
- a priori $\sup(\text{D}) \leq \inf(\text{MK})$

$$\int_{\mathcal{X}} \varphi d\mu_0 + \int_{\mathcal{Y}} \psi d\mu_1 = \int_{\mathcal{X} \times \mathcal{Y}} \varphi \oplus \psi d\pi \leq \int_{\mathcal{X} \times \mathcal{Y}} c d\pi$$

Theorem (informal)

- ① for any optimal $(\pi; (\varphi, \psi))$, $\varphi \oplus \psi = c$ on $\text{supp}(\pi)$
- ② $\inf(\text{MK}) = \sup(\text{D})$

Duality

$$\int_{\mathcal{X}} \varphi(x) \mu_0(dx) + \int_{\mathcal{Y}} \psi(y) \mu_1(dy) \rightarrow \max; \quad \varphi, \psi : \varphi \oplus \psi \leq c \quad (\text{D})$$

- Improving (φ, ψ)

- $\varphi(x) + \psi(y) \leq c(x, y)$
- $\psi(y) \leq c(x, y) - \varphi(x)$
improved by: $\varphi^c(y) := \inf_x \{c(x, y) - \varphi(x)\}$
- $\varphi(x) \leq c(x, y) - \varphi^c(y)$
improved by: $\varphi^{cc}(x) := \inf_y \{c(x, y) - \varphi^c(y)\}$
- $\varphi^{cc}(x) + \varphi^c(y) \leq c(x, y)$ is our best: $\varphi^{ccc} = \varphi^c$

Taking $\varphi = \varphi^{cc}$, $\psi = \varphi^c$ saturates $\varphi^{cc} \oplus \varphi^c \leq c$

This means if (φ, ψ) solves (D), then

- $(\varphi^{cc}, \varphi^c)$ also solves (D)
- $\varphi \oplus \psi = \varphi^{cc} \oplus \varphi^c = c$, π -a.e., for any solution π of (MK)

Duality ($c = d$)

- (\mathcal{X}, d) metric space. $\mathcal{X} = \mathcal{Y}$
 - ▶ φ^d is a 1-Lipschitz function
 - ★ $\varphi^d = \inf_x \{d(x, \cdot) - \varphi(x)\}$: inf of 1-Lipschitz functions
 - ▶ $\varphi^{dd} = -\varphi^d$
 - ★ $\varphi^{dd}(x) + \varphi^d(x) \leq d(x, x) = 0$, i.e. $-\varphi^d \geq \varphi^{dd}$
 - ★ φ^d is 1-Lip $\Rightarrow -\varphi^d(x) \leq d(x, y) - \varphi^d(y), \forall y \Rightarrow -\varphi^d \leq \varphi^{dd}$
 - ▶ $\varphi^{dd}(x) + \varphi^d(y) \leq c(x, y) \iff \varphi^d(y) - \varphi^d(x) \leq d(x, y)$

Theorem (Kantorovich-Rubinstein)

For all $\mu_0, \mu_1 \in \mathcal{P}_1(\mathcal{X})$, $W_1(\mu_0, \mu_1) = \sup_{\varphi \in \text{Lip}_1} \int_{\mathcal{X}} \varphi \, d(\mu_1 - \mu_0)$

Quadratic transport on $\mathcal{X} = \mathbb{R}^n$

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |y - x|^2 \pi(dx dy) \rightarrow \min; \quad \pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n) : \pi_0 = \mu_0, \pi_1 = \mu_1$$

Theorem ($\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$, $c(x, y) = |y - x|^2$)

Suppose $\mu_0(dx) \ll dx$, $\mu_0, \mu_1 \in \mathcal{P}_2(\mathcal{X})$. Then,
 π^\top is the unique solution of (MK) where $T = \nabla \theta$, μ_0 -a.e. with θ convex

Cuesta-Albertos-Matran (89), Rüschorndorf-Rachev (90), Brenier (91)

- sketch of proof. $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$, $c(x, y) = |y - x|^2/2$
 - ▶ $\varphi^{cc}(x) + \varphi^c(y) \leq c(x, y) = |x|^2/2 + |y|^2/2 - x \cdot y$
 - ▶ $\theta(x) := |x|^2/2 - \varphi^{cc}(x)$, $\theta^*(y) := |y|^2/2 - \varphi^c(y)$
 - ▶ $x \cdot y \leq \theta(x) + \theta^*(y)$
 - ▶ $\theta^*(y) = \sup_x \{x \cdot y - \theta(x)\}$, $\theta^{**}(x) := \sup_y \{x \cdot y - \theta^*(y)\} = \theta(x)$
 - ▶ $\theta = \theta^{**}$ is convex and differentiable a.e.
 - ▶ $\varphi^{cc}(x) + \varphi^c(y) = c(x, y) \iff \theta(x) + \theta^*(y) = x \cdot y \iff y \in \partial \theta(x)$
 - ▶ if $\mu_0(dx) \ll dx$, θ is differentiable μ_0 -a.e. Hence, $y = \nabla \theta(x)$, π -a.e.

Quadratic transport on $\mathcal{X} = \mathbb{R}^n$

- $\mathcal{X} = \mathbb{R}$
 - ▶ θ convex, $T = \theta'$ is increasing

Monotone rearrangement

$$T = F_1^{-1} \circ F_0$$

$$F_0(x) := \mu_0((-\infty, x]), \quad F_1(y) := \mu_1((-\infty, y])$$

- $\mathcal{X} = \mathbb{R}^n$
 - ▶ $T_{\#}\mu_0 = \mu_1 \iff \mu_0(x) = \mu_1(T(x)) |\det \nabla T(x)|, \forall x$
 - ▶ $T = \nabla \theta$

Monge-Ampère equation

$$\mu_1(\nabla \theta) \det \text{Hess } \theta = \mu_0, \quad \theta \text{ convex}$$

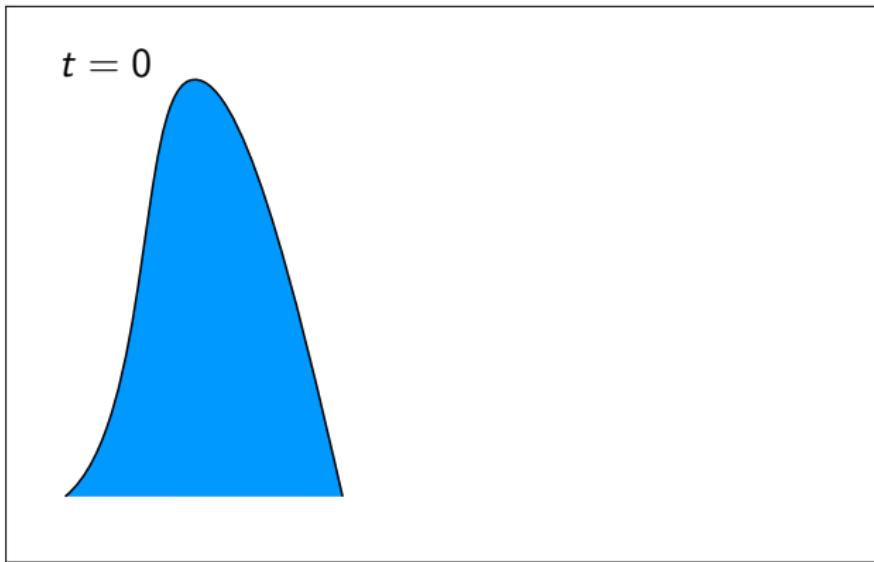
Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



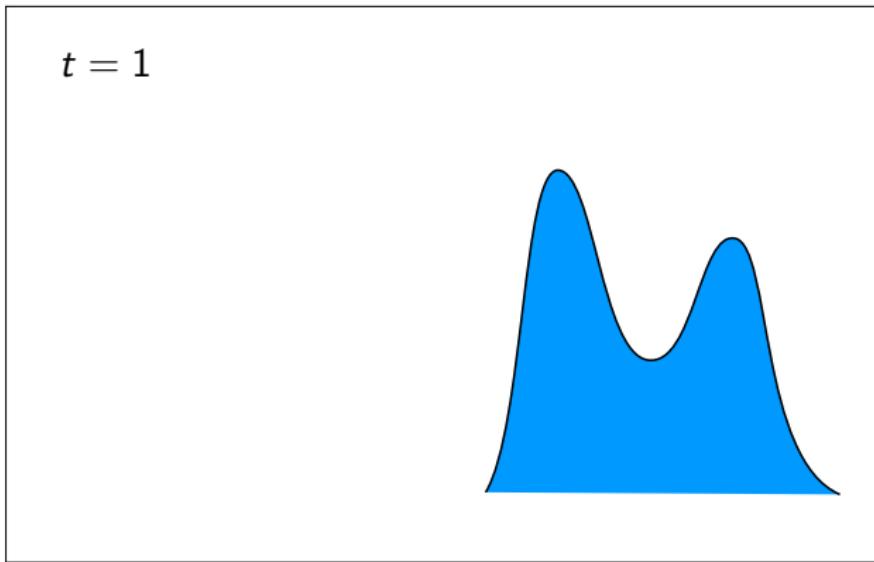
Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



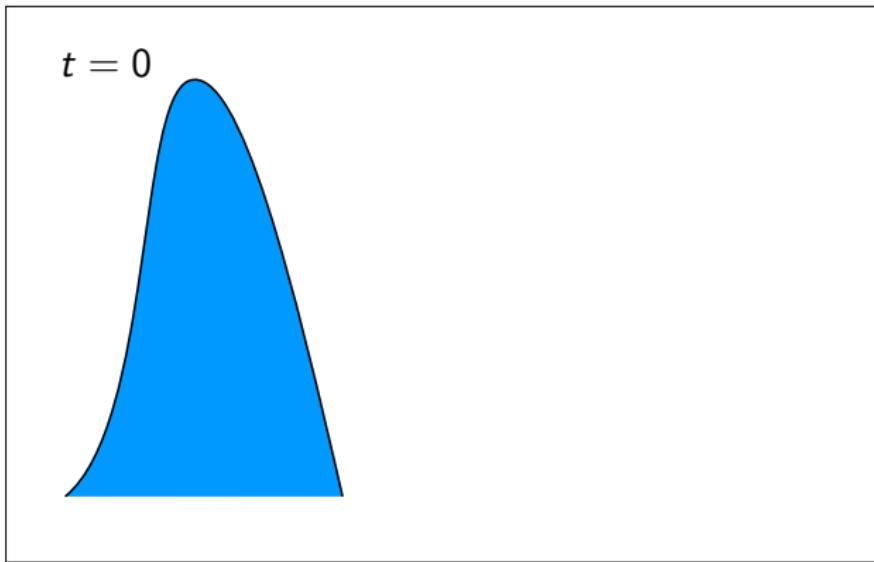
Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



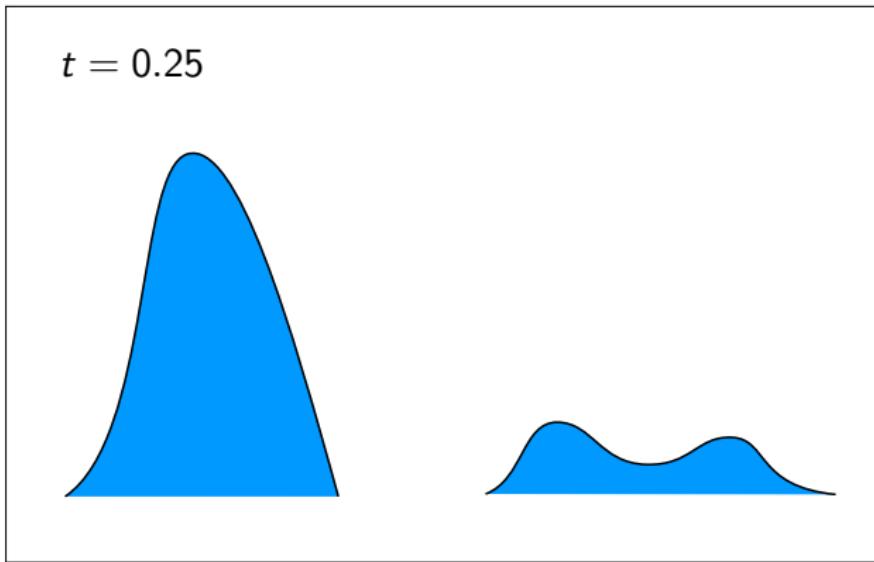
Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



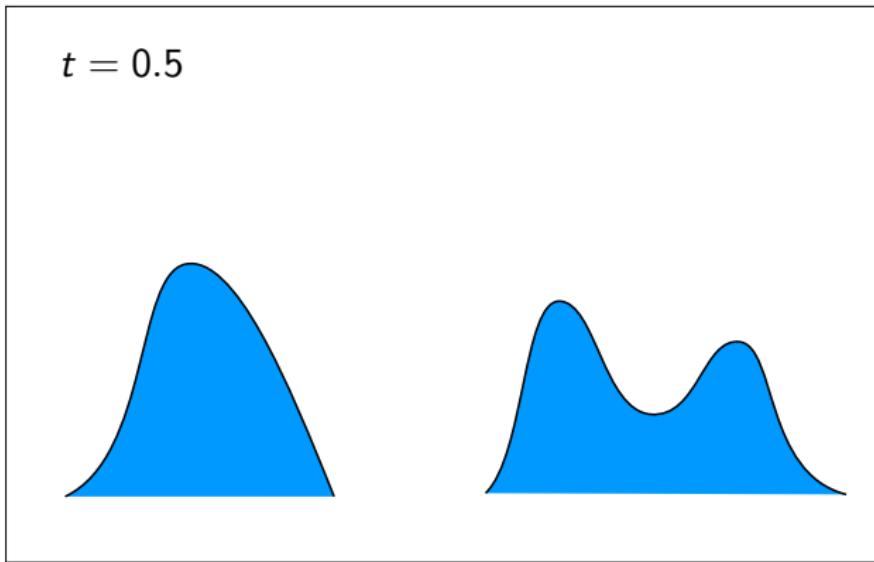
Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



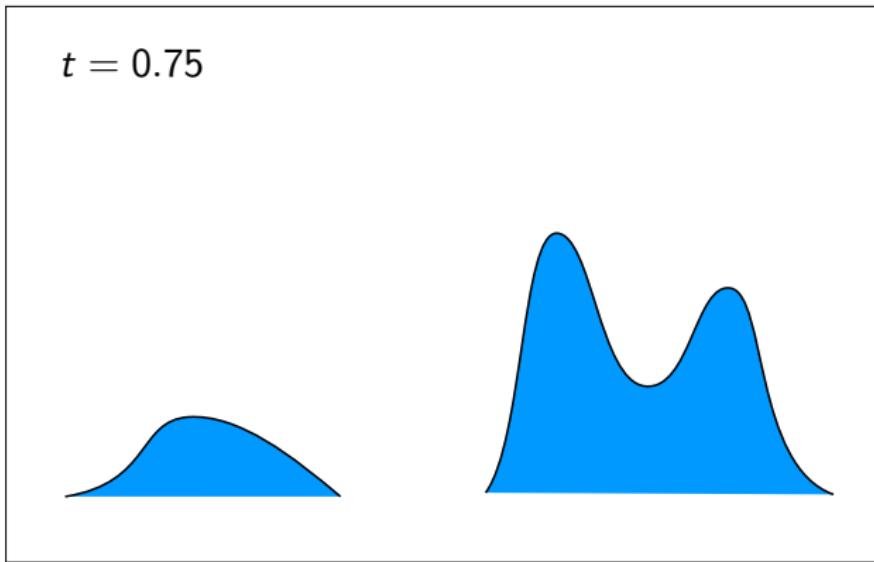
Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



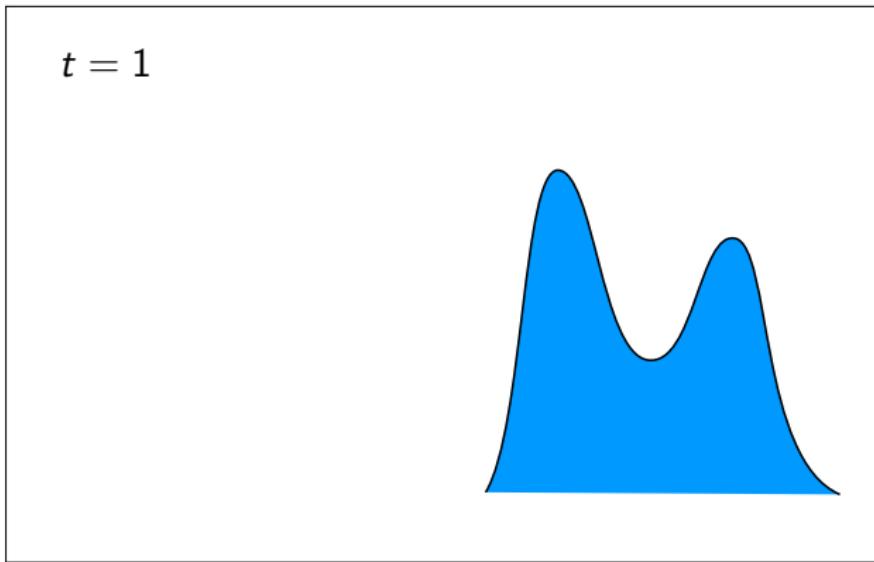
Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



Interpolations in $\mathcal{P}(\mathcal{X})$

- Standard affine interpolation between μ_0 and μ_1
 $\mu_t^{\text{aff}} := (1 - t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), 0 \leq t \leq 1$



Interpolations in $\mathcal{P}(\mathcal{X})$

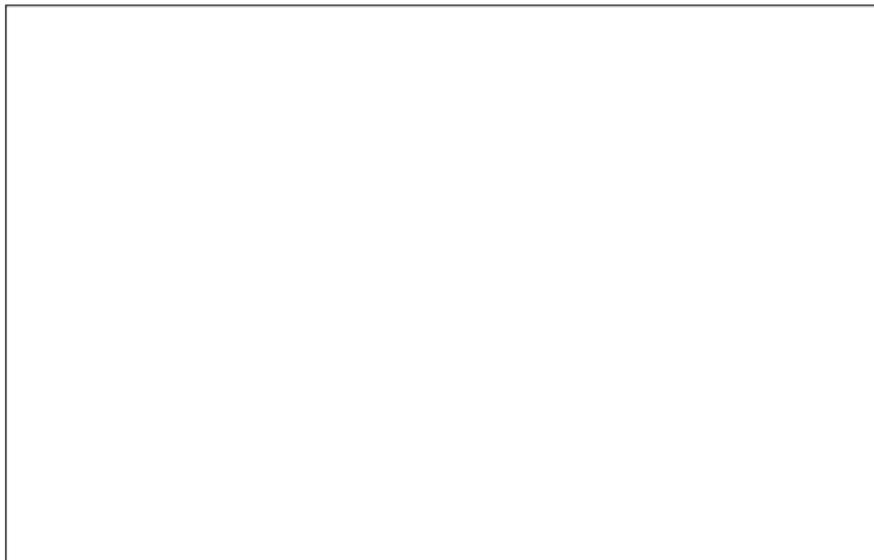
Affine interpolations require mass transference with infinite speed



- Denial of the geometry of \mathcal{X}
- We need interpolations built upon *trans*-portation, not *tele*-portation

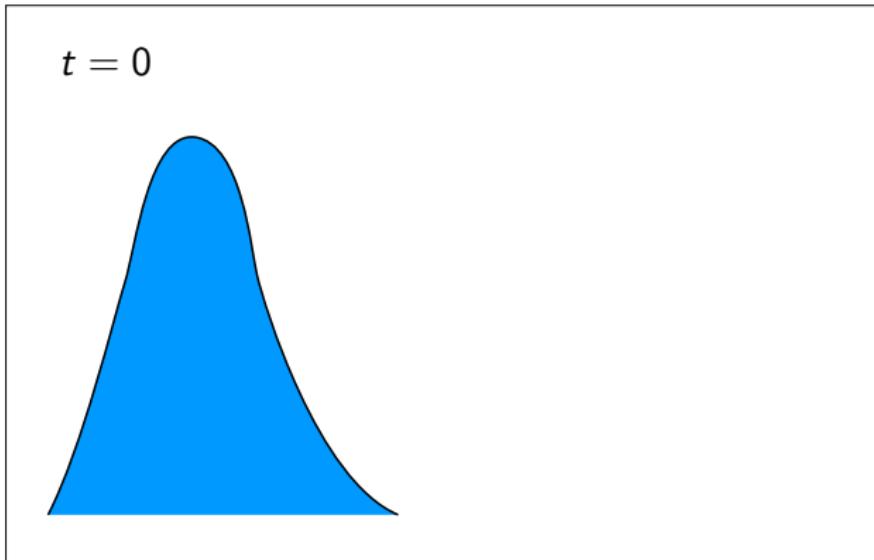
Interpolations in $\mathcal{P}(\mathcal{X})$

- We seek interpolations of this type



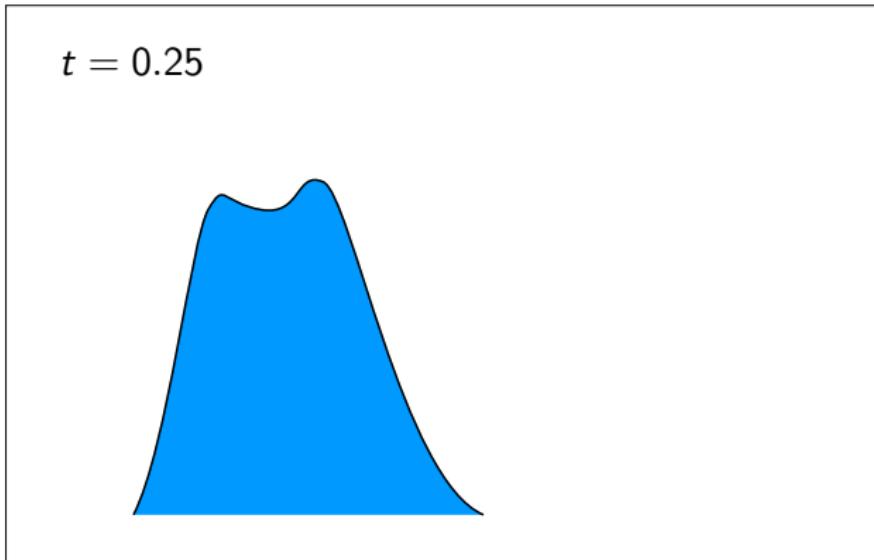
Interpolations in $\mathcal{P}(\mathcal{X})$

- We seek interpolations of this type



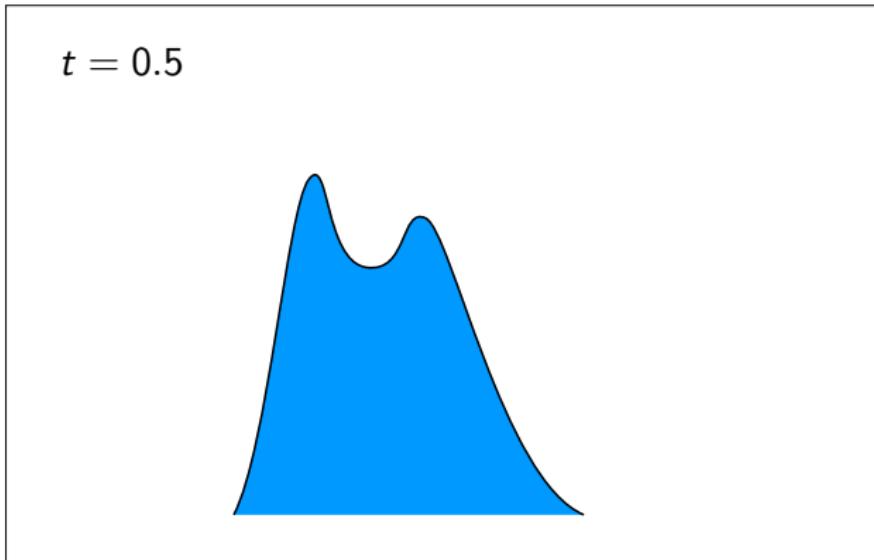
Interpolations in $\mathcal{P}(\mathcal{X})$

- We seek interpolations of this type



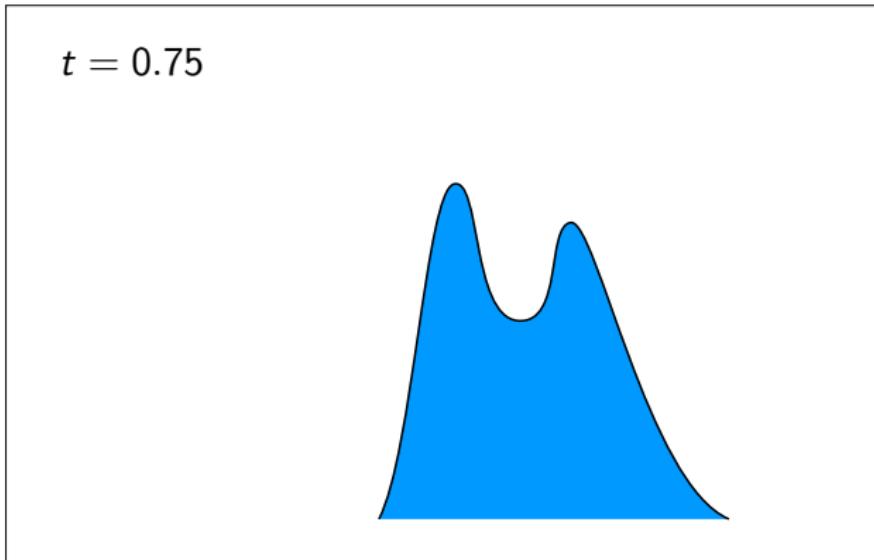
Interpolations in $\mathcal{P}(\mathcal{X})$

- We seek interpolations of this type



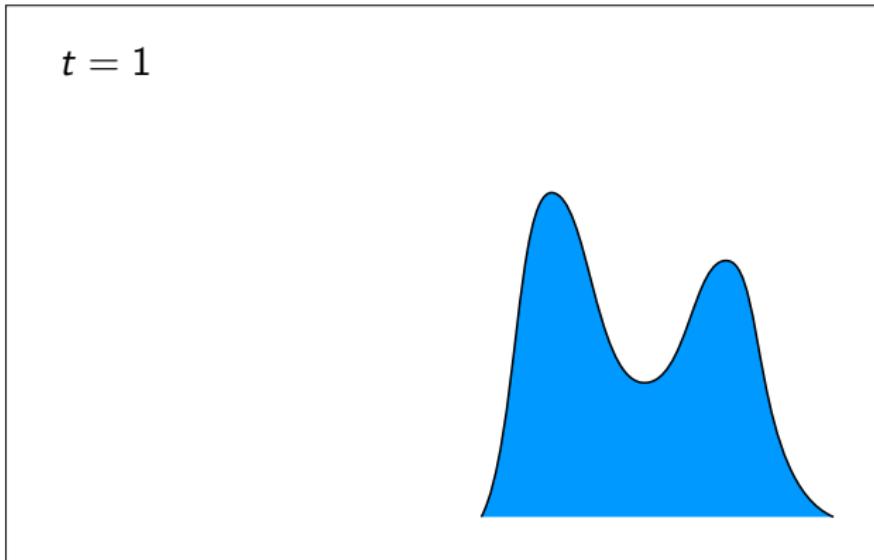
Interpolations in $\mathcal{P}(\mathcal{X})$

- We seek interpolations of this type



Interpolations in $\mathcal{P}(\mathcal{X})$

- We seek interpolations of this type



Notation

- $\mathcal{X} = \mathcal{Y}$
- $\Omega = \{\text{paths}\} \subset \mathcal{X}^{[0,1]}$
- $\omega = (\omega_t)_{0 \leq t \leq 1} \in \Omega$
- $X_t : \omega \in \Omega \mapsto \omega_t \in \mathcal{X}, \quad 0 \leq t \leq 1 \quad (\text{canonical process})$
- $P \in \mathcal{P}(\Omega)$
- $P_t := (X_t)_\# P \in \mathcal{P}(\mathcal{X})$
- $P_{st} = (X_s, X_t)_\# P \in \mathcal{P}(\mathcal{X}^2)$
- P_0 : initial marginal, P_1 : final marginal, P_{01} : endpoint marginal
- $P^{xy} = P(\cdot \mid X_0 = x, X_1 = y) \in \mathcal{P}(\Omega)$: bridge

Disintegration formula

$$P(\cdot) = \int_{\mathcal{X}^2} P^{xy}(\cdot) P_{01}(dxdy) \in \mathcal{P}(\Omega)$$

Dynamical quadratic transport on $\mathcal{X} = \mathbb{R}^n$

Geodesic equality

$$|y - x|^2 = \inf_{\omega} \int_0^1 |\dot{\omega}_t|^2 dt; \quad \omega = (\omega_t)_{0 \leq t \leq 1} : \omega_0 = x, \omega_1 = y$$

Minimizer: constant speed geodesic $\gamma_t^{xy} = (1-t)x + ty$, $0 \leq t \leq 1$

Dynamical Monge-Kantorovich problem

$$E_P \int_0^1 |\dot{X}_t|^2 dt \rightarrow \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1$$

Theorem

- ① \widehat{P} solution iff $\widehat{P}(\cdot) = \int_{\mathcal{X}^2} \delta_{\gamma^{xy}}(\cdot) \widehat{\pi}(dxdy)$ where $\widehat{\pi}$ solves (MK)₂
- ② $E_{\widehat{P}} \int_0^1 |\dot{X}_t|^2 dt = W_2^2(\mu_0, \mu_1)$

Proof. $E_P \int_0^1 |\dot{X}_t|^2 dt = \int_{\mathcal{X}^2} [E_{P^{xy}} \int_0^1 |\dot{X}_t|^2 dt] P_{01}(dxdy)$
 $\geq \int_{\mathcal{X}^2} |y - x|^2 P_{01}(dxdy) \geq \int_{\mathcal{X}^2} |y - x|^2 \widehat{\pi}(dxdy) = W_2^2(\mu_0, \mu_1)$ \square

Displacement interpolation

When $\mu_0(dx) \ll dx$, there is a *unique* solution

$$\widehat{P}(\cdot) = \int_{\mathcal{X}} \delta_{\gamma^x, \nabla \theta(x)}(\cdot) \mu_0(dx) = \int_{\mathcal{X}^2} \delta_{\gamma^{xy}}(\cdot) \widehat{\pi}(dxdy)$$

Displacement interpolation (McCann)

$$[\mu_0, \mu_1] = (\mu_t)_{0 \leq t \leq 1} \quad \text{where} \quad \mu_t := \widehat{P}_t \in \mathcal{P}(\mathcal{X})$$

$$\mu_t = \widehat{P}_t = \int_{\mathcal{X}} \delta_{\gamma_t^x, \nabla \theta(x)}(\cdot) \mu_0(dx) = \int_{\mathcal{X}^2} \delta_{\gamma_t^{xy}}(\cdot) \widehat{\pi}(dxdy) \in \mathcal{P}(\mathcal{X})$$

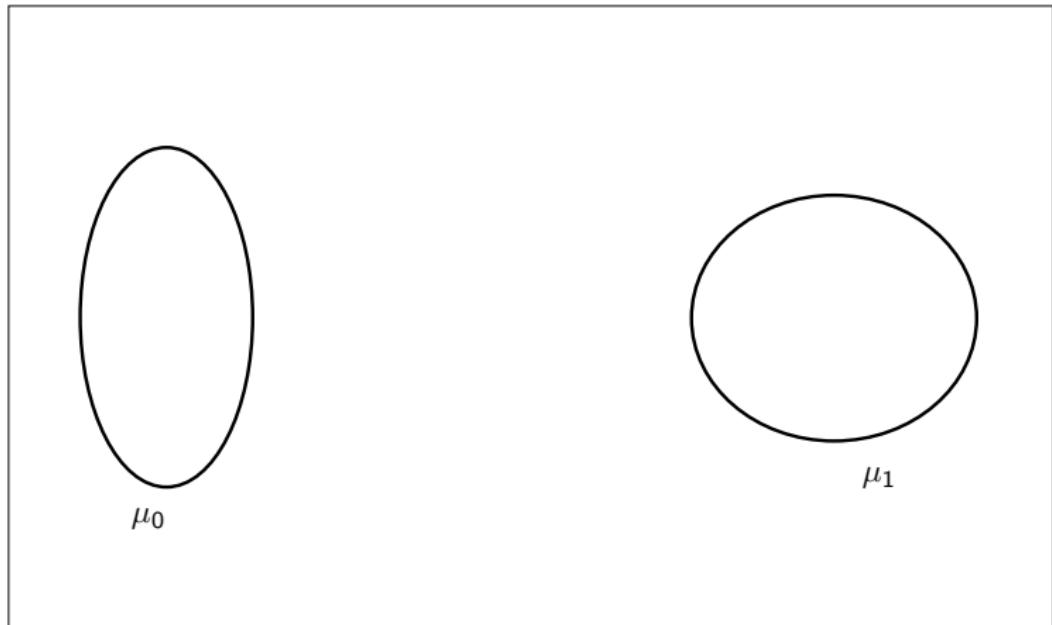
Proposition

For all $0 \leq s, t \leq 1$, \widehat{P}_{st} is an optimal coupling of (μ_s, μ_t)

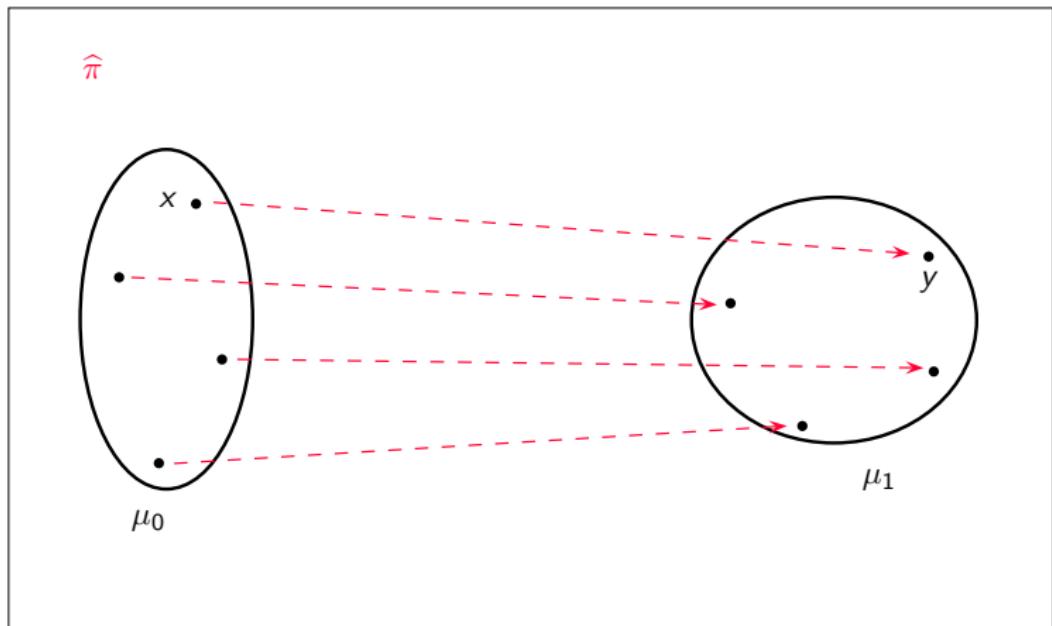
Proof. Otherwise, \widehat{P} wouldn't be optimal

□

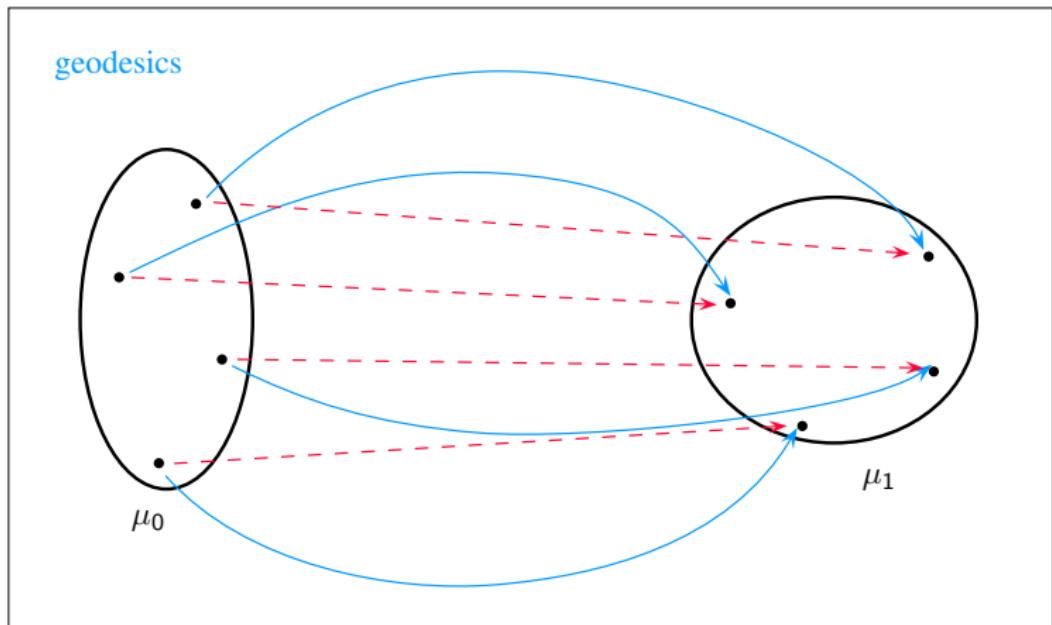
Displacement interpolation



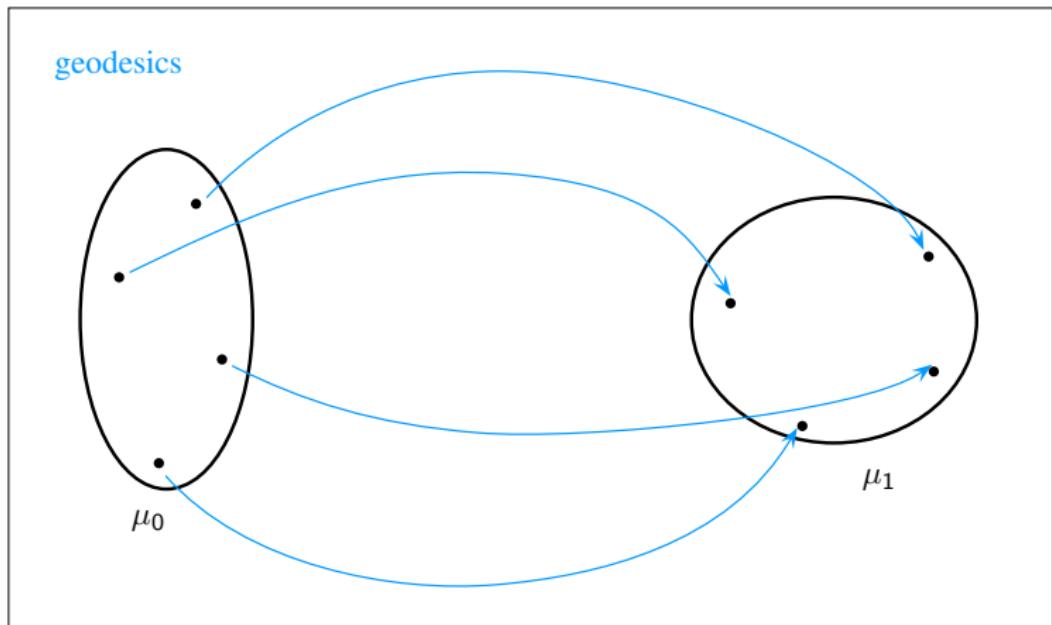
Displacement interpolation



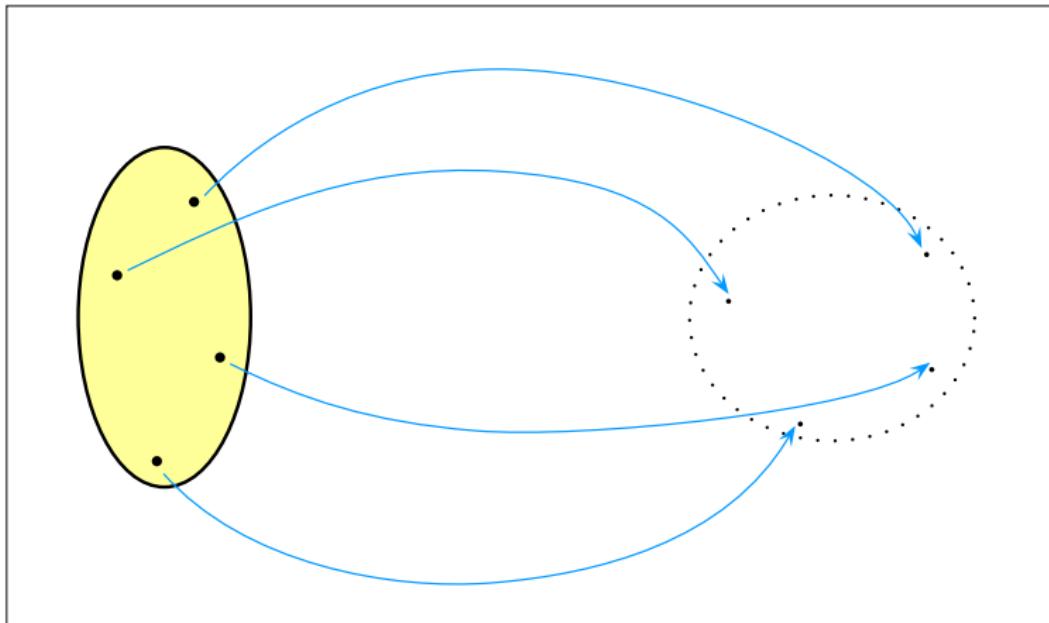
Displacement interpolation



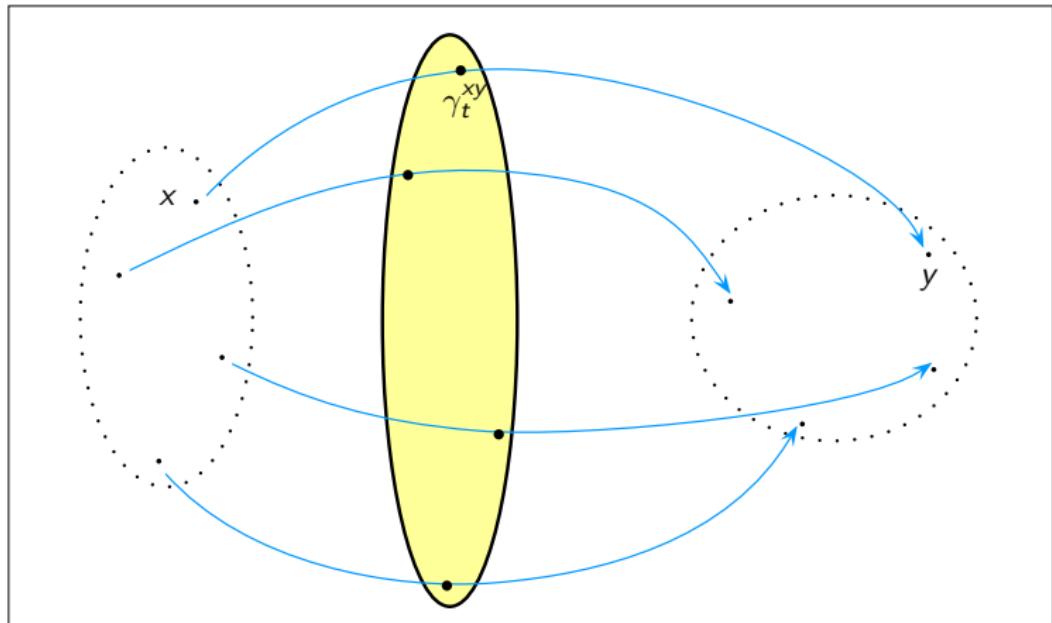
Displacement interpolation



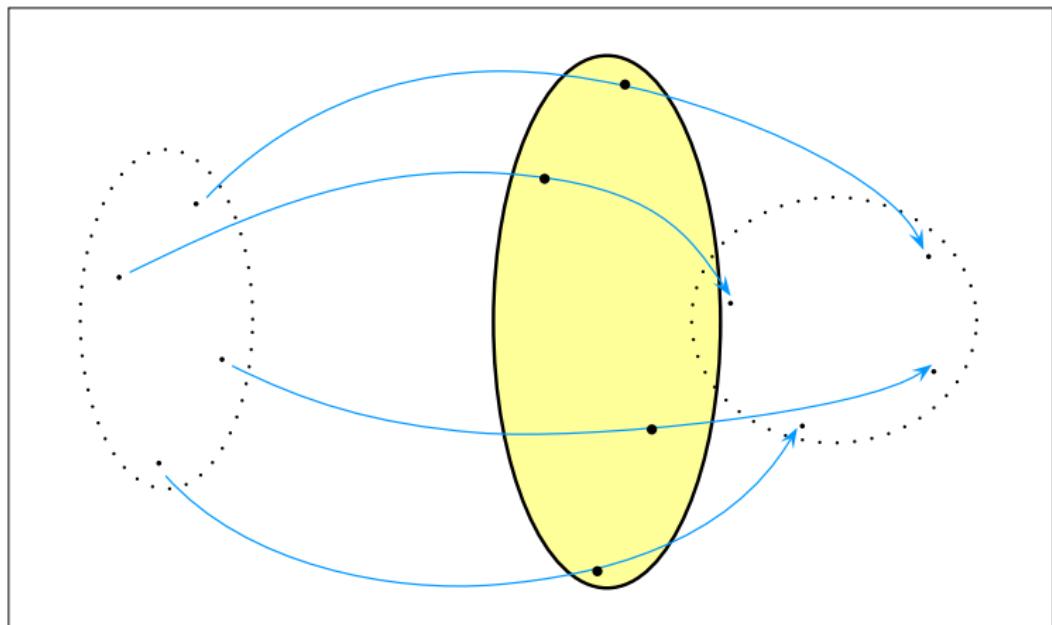
Displacement interpolation



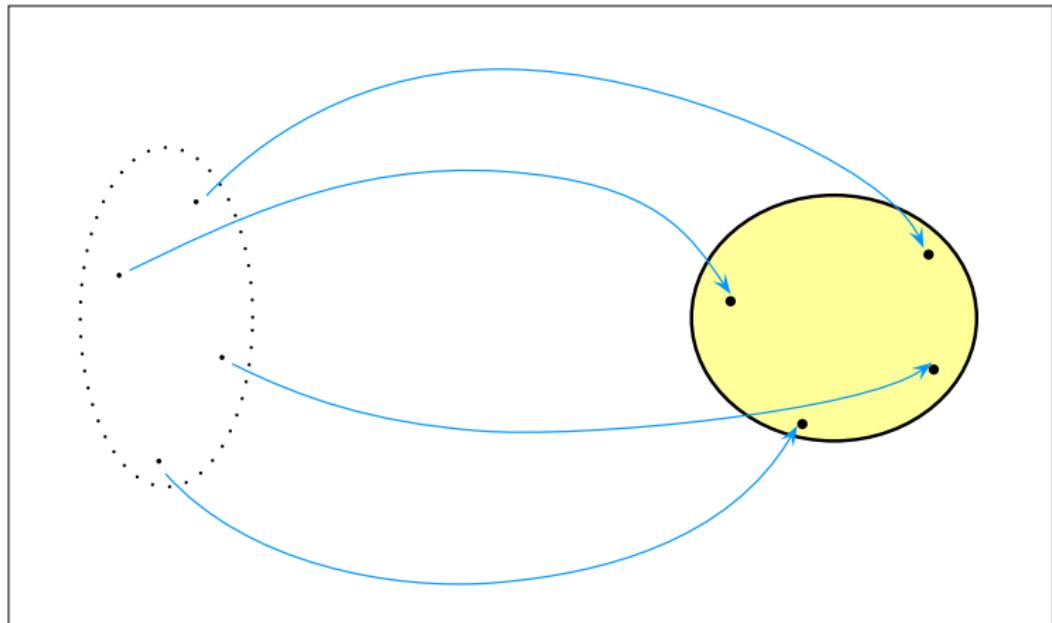
Displacement interpolation



Displacement interpolation



Displacement interpolation



Displacement interpolation, $\mathcal{X} = (M, g)$

Forward velocity field: $\nabla \psi$

$$\begin{cases} \partial_t \psi_t + |\nabla \psi_t|^2/2 = 0; & \leftarrow \psi_1 = \psi \\ \frac{\partial}{\partial t} \mu_t + \nabla \cdot (\mu_t \nabla \psi_t) = 0; & \mu_0 \rightarrow \end{cases}$$

- Sketch of proof.

- $\begin{cases} \theta = |\cdot|^2/2 - \varphi \\ \theta^* = |\cdot|^2/2 - \psi \end{cases}, \quad \begin{cases} y = \nabla \theta(x) \\ x = \nabla \theta^*(y) \end{cases}, \quad \widehat{\pi}\text{-a.e.}$
- $\dot{\gamma}_t^{x,y} = y - x = y - \nabla \theta^*(y) = \nabla \psi(y), \forall t, \quad \widehat{\pi}\text{-a.e.}$
- \widehat{P}_{t1} optimal coupling of (μ_t, μ_1) implies: $\dot{\gamma}_t^{xy} = \nabla \psi_t(\gamma_t^{xy})$ for some ψ_t
- $0 = \ddot{\gamma}_t^{xy} = \nabla (\partial_t \psi_t + |\nabla \psi_t|^2/2)$ implies $\partial_t \psi_t + |\nabla \psi_t|^2/2 = c(t)$
- $\psi_t(z) \leftrightarrow \psi_t(z) - C(t)$ leads to: $\partial_t \psi_t + |\nabla \psi_t|^2/2 = 0$ \square

Backward velocity field: $\nabla \varphi$

$$\begin{cases} -\partial_t \varphi_t + |\nabla \varphi_t|^2/2 = 0; & \varphi_0 = \varphi \rightarrow \\ -\frac{\partial}{\partial t} \mu_t + \nabla \cdot (\mu_t \nabla \varphi_t) = 0; & \leftarrow \mu_1 \end{cases}$$

Displacement interpolation, $\mathcal{X} = (M, g)$

Benamou-Brenier formula

For all $0 \leq t \leq 1$, $\langle |\nabla \psi_t|^2, \mu_t \rangle = W_2^2(\mu_0, \mu_1)$ and

$$W_2^2(\mu_0, \mu_1) = \inf_{(\nu, \nu)} \int_0^1 \langle |\nu_t|^2, \nu_t \rangle dt = \int_0^1 \langle |\nabla \psi_t|^2, \mu_t \rangle dt$$

where (ν, ν) satisfies $\begin{cases} \frac{\partial}{\partial t} \nu_t + \nabla \cdot (\nu_t \nu_t) = 0 \\ \nu_0 = \mu_0, \quad \nu_1 = \mu_1. \end{cases}$

- ▶ $\mu = [\mu_0, \mu_1]$ is a constant speed geodesic
- ▶ The inf is attained at $(\nu, \nu) = (\nabla \psi, \mu)$
- Sketch of proof.
 - ▶ For all $u : \mathbb{R}^n \rightarrow \mathbb{R}$ and $0 \leq t \leq 1$,
$$\begin{aligned} \langle u(\nabla \psi_t), \mu_t \rangle &= E_{\widehat{P}} u(\nabla \psi_t) = E_{\widehat{P}} u(\dot{X}_t) \\ &= E_{\widehat{P}} u(X_1 - X_0) = \int_{\mathcal{X}^2} u(y - x) \pi(dx dy) \end{aligned}$$
 - ▶ $W_2^2 = \inf_P \int_0^1 |\dot{X}_t|^2 dt = E_{\widehat{P}} \int_0^1 |\dot{X}_t|^2 dt = \int_0^1 \langle |\nabla \psi_t|^2, \mu_t \rangle dt$ □

Displacement interpolation, $\mathcal{X} = (M, g)$

Otto's formal calculus

$(\mathcal{P}_2(\mathcal{X}), W_2)$ looks like a Riemannian manifold

Its constant speed geodesics are the displacement interpolations

- $H(p|r) := \int \log(dp/dr) dp$, relative entropy
 p : probability measure, r : reference σ -finite measure
- vol : volume measure of the Riemannian manifold \mathcal{X}

Theorem (Sturm-von Renesse, 2004)

The following are equivalent

- $\text{Ric} \geq K$
- for any $[\mu_0, \mu_1]$, $t \in [0, 1] \mapsto H(\mu_t|\text{vol})$ is K -convex on $(\mathcal{P}_2(\mathcal{X}), W_2)$
- $H(\mu_t|m) \leq (1-t)H(\mu_0|m) + tH(\mu_1|m)$
 $- KW_2^2(\mu_0, \mu_1)t(1-t)/2$, $0 \leq t \leq 1$
- non-smooth formulation of $\frac{d^2}{dt^2} H(\mu_t|m) \geq KW_2^2(\mu_0, \mu_1)$

Analytic consequences

- $m = e^{-V} \text{vol}$

Entropic convexity

The following are equivalent

- $\text{Hess } V + \text{Ric} \geq K$
- for any $[\mu_0, \mu_1]$, $t \in [0, 1] \mapsto H(\mu_t|m)$ is K -convex on $(\mathcal{P}_2(\mathcal{X}), W_2)$

When $K > 0$:

- m satisfies an entropy-energy inequality (logarithmic Sobolev)
- exponential convergence to m as $t \rightarrow \infty$ of heat flows
- m satisfies an transport-entropy inequality (Talagrand)
- concentration of the measure m
- isoperimetric inequality

$$\text{Extension: } (M, d_g, \text{vol}_g) \rightarrow (\mathcal{X}, d, m)$$

- (\mathcal{X}, d, m) metric measure space
- d allows for addressing $(\text{MK})_2$
good notion of displacement interpolation on $(\mathcal{P}_2(\mathcal{X}), W_2)$
- m allows for defining $H(\cdot|m)$

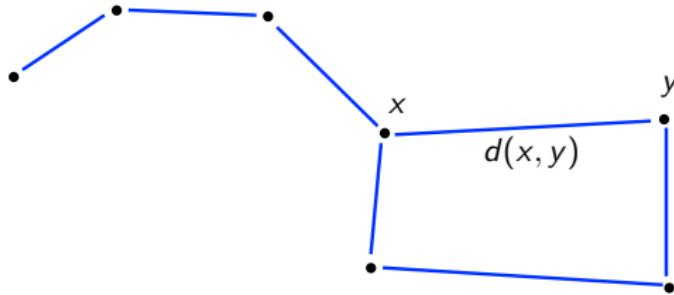
Definition of curvature lower bound (Lott-Sturm-Villani)

Framework: (\mathcal{X}, d) is a non-branching geodesic space
 (\mathcal{X}, d, m) has curvature lower bound K if along any displacement interpolation $[\mu_0, \mu_1]$, $t \mapsto H(\mu_t|m)$ is K -convex on $(\mathcal{P}_2(\mathcal{X}), W_2)$.

- (works with Gromov-Hausdorff limits of Riemannian manifolds
- (fails with graphs
 - ▶ no *constant speed* geodesics
 - ▶ branching

Regular geodesic on a graph

- non-directed metric graph (\mathcal{X}, \sim, d)



$x \sim y$ means that (x, y) is an edge

Problem

How to define a constant speed geodesic on (\mathcal{X}, \sim) ?

- one *must* introduce a random walk on (\mathcal{X}, \sim)
- seek a regular geodesic $[\mu_0, \mu_1] = (\mu_t)_{0 \leq t \leq 1}$ on $\mathcal{P}(\mathcal{X})$
- it will allow $\frac{d^2}{dt^2} H(\mu_t | m)$

Regular geodesic on a graph

- length $\ell(\omega) := \sum_{0 \leq t \leq 1} \mathbf{1}_{\{\omega_{t-} \neq \omega_t\}} d(\omega_{t-}, \omega_t)$
- intrinsic distance $d(x, y) = \inf \{\ell(\omega) : \omega \in \Omega, \omega_0 = x, \omega_1 = y\}$

Geodesics

$$\Gamma^{xy} := \{\omega \in \Omega; \omega_0 = x, \omega_1 = y, \ell(\omega) = d(x, y)\}$$

$$\Gamma := \bigcup_{x,y} \Gamma^{xy}$$

Lazy random walks

- to recover d :
 - ▶ slow down the walk
 - ▶ condition at $t = 0$ and $t = 1$
- reference walk: $R \in \mathcal{P}(\Omega)$ with jump kernel
$$J_x(dy) = \sum_{y:y \sim x} J_x(y) \delta_y$$

Lazy random walks R^k , $k \rightarrow \infty$

$$J_x^k(dy) = \sum_{y:y \sim x} k^{-d(x,y)} J_x(y) \delta_y$$

Convergence of bridges

$$\lim_{k \rightarrow \infty} R^{k,xy} = G^{xy} \in \mathcal{P}(\Gamma^{xy})$$

- $G := \mathbf{1}_\Gamma e^{\int_0^1 J_{X_t}(\mathcal{X}) dt} R$
- proof based on entropy minimization

The metric measure graph (\mathcal{X}, d, m)

- take a measure $m = (m_x)_{x \in \mathcal{X}}$.
- choose $J_x(y) = s(x, y) \sqrt{\frac{m_y}{m_x}}$
- R^k is ***m-reversible***, for any $k \geq 1$

(\mathcal{X}, d, m)

The asymptotic behaviour of $(R^k)_{k \geq 1}$ allows for seeing (\mathcal{X}, d, m) as a metric measure space

- interest
 - ▶ concentration of the measure m
 - ▶ isoperimetry

Slowing down

Dynamical Schrödinger problem

$$H(P|R^k)/\log k \rightarrow \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1 \quad (\text{S}_{\text{dyn}}^k)$$

Dynamical Monge-Kantorovich problem

$$\int_{\Omega} \ell \, dP \rightarrow \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1 \quad (\text{MK}_{\text{dyn}})$$

Theorem

- Γ - $\lim_{k \rightarrow \infty} (\text{S}_{\text{dyn}}^k) = (\text{MK}_{\text{dyn}})$
- $\lim_{k \rightarrow \infty} \inf(\text{S}_{\text{dyn}}^k) = \inf(\text{MK}_{\text{dyn}})$
- $\lim_{k \rightarrow \infty} P^k = P^*$: singled out solution of (MK_{dyn})

Slowing down

Schrödinger problem

$$H(\pi|R_{01}^k)/\log k \rightarrow \min; \quad \pi \in \mathcal{P}(\mathcal{X}^2) : \pi_0 = \mu_0, \pi_1 = \mu_1 \quad (S^k)$$

Monge-Kantorovich problem

$$\int_{\mathcal{X}^2} d d\pi \rightarrow \min; \quad \pi \in \mathcal{P}(\mathcal{X}^2) : \pi_0 = \mu_0, \pi_1 = \mu_1 \quad (\text{MK})$$

Theorem

- Γ - $\lim_{k \rightarrow \infty} (S^k) = (\text{MK})$
- $\lim_{k \rightarrow \infty} \inf(S^k) = \inf(\text{MK}) := W_1(\mu_0, \mu_1) = \inf(\text{MK}_{\text{dyn}})$
- $\lim_{k \rightarrow \infty} \pi^k = \pi^*$: singled out solution of (MK)

Slowing down

Convergence schema

$$\begin{array}{ccc} P^k(d\omega) & = & \int_{\mathcal{X}^2} R^k(d\omega \mid X_0 = x, X_1 = y) \pi^k(dx dy) \\ \downarrow & & \downarrow \\ P^*(d\omega) & = & \int_{\mathcal{X}^2} G^{xy}(d\omega) \pi^*(dx dy) \end{array}$$

Theorem

P^* is Markov and $\text{supp } P^* \subset \Gamma$

Entropic interpolations converge to displacement interpolations

$$\begin{array}{ccc} \mu_t^k(dz) & = & \int_{\mathcal{X}^2} R_t^k(dz \mid X_0 = x, X_1 = y) \pi^k(dx dy) \\ \downarrow & & \downarrow \\ \mu_t^*(dz) & = & \int_{\mathcal{X}^2} G_t^{xy}(dz) \pi^*(dx dy) \end{array}$$

Slowing down

Recall

- $\mu_t(dz) = \int_{\mathcal{X}^2} \delta_{\gamma_t^{xy}}(dz) \widehat{\pi}(dxdy)$
- $\mu_t^*(dz) = \int_{\mathcal{X}^2} G_t^{xy}(dz) \pi^*(dxdy)$

From manifolds to graphs

$$\begin{array}{ccc} \text{deterministic clock} & \rightarrow & \text{random clock} \\ \delta_{\gamma^{xy}} & \rightarrow & G^{xy} \end{array}$$

Example:

- when $J_x(\mathcal{X}) = \lambda, \forall x$ and $d(x, y) = 1, \forall x \sim y$
 G^{xy} is a mixture of Poisson bridges
- for instance, simple random walk: $J_x = n_x^{-1} \sum_{y:y \sim x} \delta_y, m = \sum_x n_x \delta_x$

Benamou-Brenier formula

$\inf_{(\nu, K)}$

$$\int_{[0,1]} \left(\sum_{x,y:x \sim y} d(x,y) K_{t,x}(y) \nu_t(x) \right) dt$$

where (ν, K) is such that

$$\begin{cases} \partial_t \nu_t(x) - \sum_{y:y \sim x} \{ \nu_t(y) K_{t,y}(x) - \nu_t(x) K_{t,x}(y) \} = 0 \\ \nu_0 = \mu_0, \nu_1 = \mu_1 \end{cases}$$

Theorem

- The solution is the displacement interpolation:
 $\nu = \mu^*$ and $K = J^{\mu^*}$
- $W_1(\mu_0, \mu_1) = \inf_{(\nu, j)} \int_{[0,1]} \left(\sum_{x,y:x \sim y} d(x,y) K_{t,x}(y) \nu_t(x) \right) dt$
 $= \int_{[0,1]} \left(\sum_{x,y:x \sim y} d(x,y) J_{t,x}^{\mu^*}(y) \mu_t^*(x) \right) dt$

Constant speed interpolations

$$W_1(\mu_0, \mu_1) = \int_{[0,1]} \text{speed}(\mu^*)_t dt$$

- $\text{speed}(\mu^*)_t = \sum_{x,y: x \sim y} d(x, y) J_{t,x}^{\mu^*}(y) \mu_t^*(x)$
- $\tau : [0, 1] \rightarrow [0, 1]$: change of time
- $J_t^{\mu \circ \tau} = \tau'(t) J_{\tau(t)}^\mu$
- $W_1(\mu_0, \mu_1) = \int_{[0,1]} \text{speed}(\mu^* \circ \tau)_t dt, \quad \forall \tau$

Proposition

There exists a unique change of time τ such that

$$\text{speed}(\mu^* \circ \tau)_t = W_1(\mu_0, \mu_1), \quad \forall t$$

τ depends on (μ_0, μ_1)

Entropic convexity

The dynamics of P^k and $P^* = \lim_{k \rightarrow \infty} P^k$ are known, hence one can compute explicitly

$$\frac{d^2}{dt^2} H(\mu_t^* | m)$$

Still a lot of work to do ...



... merci de votre attention