Displacement functional and absolute continuity of Wasserstein barycenters

Jianyu MA, IMT

March 14, 2024, Séminaire Images Optimisation et Probabilités

Definitions

see the blackboard ☛

Barycenters

- Notion of mean for probability measures μ on metric spaces (E, d)
- Always exist in proper spaces `(metric spaces whose bounded closed sets are compact)`

Wasserstein spaces $(\mathcal{W}(E), W)$

- Metric spaces for optimal transport between probability measures on a Polish space `(a complete and separable metric space)`
- Wasserstein spaces are Polish spaces.

Definitions

Barycenters

- Notion of mean for probability measures μ on metric spaces (E, d)
- Always exist in proper spaces `(metric spaces whose bounded closed sets are compact)`

Wasserstein spaces $(\mathcal{W}(E), W)$

- Metric spaces for optimal transport between probability measures on a Polish space `(a complete and separable metric space)`
- Wasserstein spaces are Polish spaces.

Wasserstein barycenters

Definition

Given a Polish space (E, d), the Wasserstein space $(\mathcal{W}(E), W)$ is also Polish, over which we can construct the Wasserstein space $(\mathcal{W}(\mathcal{W}(E)), \mathbb{W})$.
Barycenters $\overline{\mu}$ of measures $\mathbb{P} \in \mathcal{W}(\mathcal{W}(E))$ are called Wasserstein barycenters.

Wasserstein barycenters

Definition

Given a Polish space (E, d), the Wasserstein space $(\mathcal{W}(E), W)$ is also Polish, over which we can construct the Wasserstein space $(\mathcal{W}(\mathcal{W}(E)), \mathbb{W})$.
Barycenters $\overline{\mu}$ of measures $\mathbb{P} \in \mathcal{W}(\mathcal{W}(E))$ are called Wasserstein barycenters.

Remark

By definition, $\mathbb{P}$ is a probability measure on $\mathcal{W}(E)$, its barycenter $\overline{\mu}$ is thus a probability measure on E.

Wasserstein barycenters

Definition

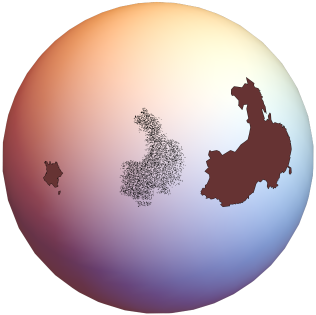
Given a Polish space (E, d), the Wasserstein space $(\mathcal{W}(E), W)$ is also Polish, over which we can construct the Wasserstein space $(\mathcal{W}(\mathcal{W}(E)), \mathbb{W})$.

Barycenters $\overline{\mu}$ of measures $\mathbb{P} \in \mathcal{W}(\mathcal{W}(E))$ are called Wasserstein barycenters.

Example (Displacement interpolation)

Consider the earth surface (E, d) with two uniform measures μ, ν supported on two regions. We simulate the barycenter of $\frac{1}{2}\delta_\mu + \frac{1}{2}\delta_\nu$ by discrete points.

🇫🇷 + 🇨🇳 $\xRightarrow{\texttt{barycenter}}$ 🦙 `(llama)`



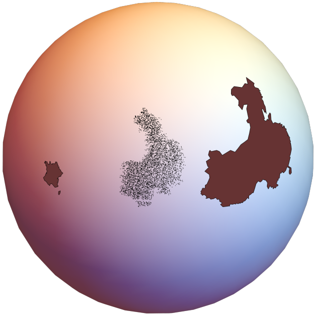
Wasserstein barycenters

Definition

Given a Polish space (E, d), the Wasserstein space $(\mathcal{W}(E), W)$ is also Polish, over which we can construct the Wasserstein space $(\mathcal{W}(\mathcal{W}(E)), \mathbb{W})$.
Barycenters $\overline{\mu}$ of measures $\mathbb{P} \in \mathcal{W}(\mathcal{W}(E))$ are called Wasserstein barycenters.

Existence [Le Gouic and Loubes, 2017]

Assuming that (E, d) is a proper space, Wasserstein barycenters in $\mathcal{W}(E)$ always exist.



Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \ldots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$ as follows.

Construction of $\overline{\mu} := B_{\#}\gamma$

1. Let $B : E^n \to E$ be a measurable map `(barycenter selection map)` sending $(x_1, x_2, \ldots, x_n)$ to a barycenter of $\sum_{i=1}^n \lambda_i \, \delta_{x_i}$.
2. Let γ be a measure `(multi-marginal optimal transport plan)` on E^n s.t.

$$\int_{E^n} c_\lambda \, \mathrm{d}\,\gamma = \inf_{\theta \in \Theta} \int_{E^n} c_\lambda \, \mathrm{d}\,\theta \quad \text{with } c_\lambda(x_1, \ldots, x_n) := \inf_{y \in E} \sum_{i=1}^n \lambda_i \, d(x_i, y)^2,$$

   where Θ is the set of measures on E^n with marginals $\mu_1, \mu_2, \ldots, \mu_n$ and $\gamma \in \Theta$.

Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \dots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \dots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$ as follows.

Construction of $\overline{\mu} := B_{\#}\gamma$

1. Let $B : E^n \to E$ be a measurable map `(barycenter selection map)` sending $(x_1, x_2, \dots, x_n)$ to a barycenter of $\sum_{i=1}^n \lambda_i \, \delta_{x_i}$.
2. Let γ be a measure `(multi-marginal optimal transport plan)` on E^n s.t.

$$\int_{E^n} c_\lambda \, \mathrm{d}\,\gamma = \inf_{\theta \in \Theta} \int_{E^n} c_\lambda \, \mathrm{d}\,\theta \quad \text{with } c_\lambda(x_1, \dots, x_n) := \inf_{y \in E} \sum_{i=1}^n \lambda_i \, d(x_i, y)^2,$$

where Θ is the set of measures on E^n with marginals $\mu_1, \mu_2, \dots, \mu_n$ and $\gamma \in \Theta$.

Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \ldots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$ as follows.

Construction of $\overline{\mu} := B_\# \gamma$

1. Let $B : E^n \to E$ be a measurable map `(barycenter selection map)` sending $(x_1, x_2, \ldots, x_n)$ to a barycenter of $\sum_{i=1}^n \lambda_i \, \delta_{x_i}$.
2. Let γ be a measure `(multi-marginal optimal transport plan)` on E^n s.t.

$$\int_{E^n} c_\lambda \, \mathrm{d}\, \gamma = \inf_{\theta \in \Theta} \int_{E^n} c_\lambda \, \mathrm{d}\, \theta \quad \text{with } c_\lambda(x_1, \ldots, x_n) := \inf_{y \in E} \sum_{i=1}^n \lambda_i \, d(x_i, y)^2,$$

where Θ is the set of measures on E^n with marginals $\mu_1, \mu_2, \ldots, \mu_n$ and $\gamma \in \Theta$.

Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \dots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \dots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$ as follows.

Why $\overline{\mu} = B_{\#}\gamma$ is a barycenter?

Notes of current step

Recall

B sends $\vec{x} = (x_1, \dots, x_n)$ to a barycenter of $\sum_{i=1}^n \lambda_i \, \delta_{x_i}$;
γ has marginals $\mu_1, \dots, \mu_n$.

$$\sum_{i=1}^n \lambda_i \, W(\mu_i, \overline{\mu})^2 \le \sum_{i=1}^n \lambda_i \int_{E^n} d(x_i, B(\vec{x}))^2 \, \mathrm{d}\,\gamma(\vec{x})$$

$$= \int_{E^n} c_\lambda(\vec{x}) \, \mathrm{d}\,\gamma(\vec{x}) \le \mathbb{E}\, c_\lambda(X_1, \dots, X_n)$$

$$\le \mathbb{E} \sum_{i=1}^n \lambda_i d(X_i, X)^2 = \sum_{i=1}^n \lambda_i \, W(\mu_i, \nu)^2$$

Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \ldots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \delta_{\mu_i}$ as follows.

Why $\overline{\mu} = B_{\#}\gamma$ is a barycenter?

Notes of current step

Recall

$c_\lambda(\vec{x})$ is the barycenter cost $\inf_{y \in E} \sum_{i=1}^n \lambda_i \, d(x_i, y)^2$

$$\sum_{i=1}^n \lambda_i W(\mu_i, \overline{\mu})^2 \leq \sum_{i=1}^n \lambda_i \int_{E^n} d(x_i, B(\vec{x}))^2 \, \mathrm{d}\,\gamma(\vec{x})$$

$$= \int_{E^n} c_\lambda(\vec{x}) \, \mathrm{d}\,\gamma(\vec{x}) \leq \mathbb{E}\, c_\lambda(X_1, \ldots, X_n)$$

$$\leq \mathbb{E} \sum_{i=1}^n \lambda_i d(X_i, X)^2 = \sum_{i=1}^n \lambda_i W(\mu_i, \nu)^2$$

Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \ldots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \delta_{\mu_i}$ as follows.

Why $\overline{\mu} = B_{\#}\gamma$ is a barycenter?

Notes of current step

Recall

γ is an optimal plan w.r.t. c_λ;
Choose r.v. X_i with law μ_i.

$$\sum_{i=1}^n \lambda_i W(\mu_i, \overline{\mu})^2 \leq \sum_{i=1}^n \lambda_i \int_{E^n} d(x_i, B(\vec{x}))^2 \,\mathrm{d}\,\gamma(\vec{x})$$

$$= \int_{E^n} c_\lambda(\vec{x}) \,\mathrm{d}\,\gamma(\vec{x}) \leq \mathbb{E}\, c_\lambda(X_1, \ldots, X_n)$$

$$\leq \mathbb{E} \sum_{i=1}^n \lambda_i d(X_i, X)^2 = \sum_{i=1}^n \lambda_i W(\mu_i, \nu)^2$$

Structure of Wasserstein barycenters

Fix a proper space (E,d) and n positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \ldots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$ as follows.

Why $\overline{\mu} = B_{\#}\gamma$ is a barycenter?

Notes of current step

Notation

X is a new r.v. with arbitrarily chosen law ν; the coupling (X_i, X) could be optimal.

$$\sum_{i=1}^n \lambda_i W(\mu_i, \overline{\mu})^2 \le \sum_{i=1}^n \lambda_i \int_{E^n} d(x_i, B(\vec{x}))^2 \,\mathrm{d}\,\gamma(\vec{x})$$

$$= \int_{E^n} c_\lambda(\vec{x}) \,\mathrm{d}\,\gamma(\vec{x}) \le \mathbb{E}\, c_\lambda(X_1, \ldots, X_n)$$

$$\le \mathbb{E} \sum_{i=1}^n \lambda_i d(X_i, X)^2 = \sum_{i=1}^n \lambda_i W(\mu_i, \nu)^2$$

Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \ldots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \delta_{\mu_i}$ as follows.

Why $\overline{\mu} = B_\# \gamma$ is a barycenter?

Notes of current step

Conclusion

Choose (X_i, X) to be optimal.

$\overline{\mu}$ is a barycenter since ν is arbitrary.

$$\sum_{i=1}^n \lambda_i W(\mu_i, \overline{\mu})^2 \leq \sum_{i=1}^n \lambda_i \int_{E^n} d(x_i, B(\vec{x}))^2 \,\mathrm{d}\,\gamma(\vec{x})$$

$$= \int_{E^n} c_\lambda(\vec{x}) \,\mathrm{d}\,\gamma(\vec{x}) \leq \mathbb{E}\, c_\lambda(X_1, \ldots, X_n)$$

$$\leq \mathbb{E} \sum_{i=1}^n \lambda_i d(X_i, X)^2 = \sum_{i=1}^n \lambda_i W(\mu_i, \nu)^2$$

Structure of Wasserstein barycenters

Fix a proper space (E, d) and n positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\sum_{i=1}^n \lambda_i = 1$. Given n measures $\mu_1, \mu_2, \ldots, \mu_n$, one can construct a barycenter $\overline{\mu}$ of $\sum_{i=1}^n \lambda_i \delta_{\mu_i}$ as follows.

Why $\overline{\mu} = B_{\#}\gamma$ is a barycenter?

Notes of current step

Corollary

Set $\nu = \overline{\mu}$; $(\mathrm{proj}_i, B)_{\#}\gamma$ is thus an optimal transport plan between μ_i and $\overline{\mu}$.

$$\sum_{i=1}^n \lambda_i W(\mu_i, \overline{\mu})^2 = \sum_{i=1}^n \lambda_i \int_{E^n} d(x_i, B(\vec{x}))^2 \,\mathrm{d}\,\gamma(\vec{x})$$

$$= \int_{E^n} c_\lambda(\vec{x}) \,\mathrm{d}\,\gamma(\vec{x}) \leq \mathbb{E}\, c_\lambda(X_1, \ldots, X_n)$$

$$\leq \mathbb{E} \sum_{i=1}^n \lambda_i d(X_i, X)^2 = \sum_{i=1}^n \lambda_i W(\mu_i, \overline{\mu})^2$$

Properties of Wasserstein barycenter

Consistency [Le Gouic and Loubes, 2017]

Let (E, d) be a proper space. Given a sequence of measures $\mathbb{P}_j \in \mathcal{W}(\mathcal{W}(E))$ with barycenters $\overline{\mu}_j$, if $\mathbb{W}(\mathbb{P}_j, \mathbb{P}) \to 0$, then $\overline{\mu}_j$ converges to a barycenter of $\mathbb{P}$ up to extracting a subsequence.

Remark

Construction for finitely many measures $+$ consistency $\implies$ general existence.

Indeed, we rely on the consistency to investigate general barycenters.

Properties of Wasserstein barycenter

Consistency [Le Gouic and Loubes, 2017]

Let (E, d) be a proper space. Given a sequence of measures $\mathbb{P}_j \in \mathcal{W}(\mathcal{W}(E))$ with barycenters $\overline{\mu}_j$, if $\mathbb{W}(\mathbb{P}_j, \mathbb{P}) \to 0$, then $\overline{\mu}_j$ converges to a barycenter of $\mathbb{P}$ up to extracting a subsequence.

Uniqueness [Kim and Pass, 2017]

Let (M, d_g) be a Riemannian manifold. If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of absolutely continuous measures, then it has a unique barycenter.

Properties of Wasserstein barycenter

Consistency [Le Gouic and Loubes, 2017]

Let (E, d) be a proper space. Given a sequence of measures $\mathbb{P}_j \in \mathcal{W}(\mathcal{W}(E))$ with barycenters $\overline{\mu}_j$, if $\mathbb{W}(\mathbb{P}_j, \mathbb{P}) \to 0$, then $\overline{\mu}_j$ converges to a barycenter of $\mathbb{P}$ up to extracting a subsequence.

Uniqueness [Kim and Pass, 2017]

Let (M, d_g) be a Riemannian manifold. If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of absolutely continuous measures, then it has a unique barycenter.

Absolute continuity [Agueh and Carlier, 2011]

Let $\mu_1, \mu_2, \ldots, \mu_n$ be n probability measures on $\mathbb{R}^m$. If μ_1 is absolutely continuous with bounded density function, then the unique barycenter of $\sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$ is also absolutely continuous.

Properties of Wasserstein barycenter

Consistency [Le Gouic and Loubes, 2017]

Let (E, d) be a proper space. Given a sequence of measures $\mathbb{P}_j \in \mathcal{W}(\mathcal{W}(E))$ with barycenters $\overline{\mu}_j$, if $\mathbb{W}(\mathbb{P}_j, \mathbb{P}) \to 0$, then $\overline{\mu}_j$ converges to a barycenter of $\mathbb{P}$ up to extracting a subsequence.

Uniqueness [Kim and Pass, 2017]

Let (M, d_g) be a Riemannian manifold. If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of absolutely continuous measures, then it has a unique barycenter.

Absolute continuity [Kim and Pass, 2017]

Let (M, d_g) be a compact Riemannian manifold. If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to a set of absolutely continuous measures with uniformly bounded density functions, then its unique barycenter is absolutely continuous.

How to prove absolute continuity

`(a.c stands for absolutely continuous)`

Absolute continuity and compactness [Kim and Pass, 2017]

Let (M, d_g) be a compact Riemannian manifold. If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to a set of a.c measures with uniformly bounded density functions, then its barycenter is a.c.

Absolute continuity and Ricci curvature bound [Ma, 2023]

Let (M, d_g) be a complete Riemannian manifold with a lower Ricci curvature bound.
If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of a.c measures, then its barycenter $\overline{\mu}$ is a.c.

How to prove absolute continuity

Absolute continuity and Ricci curvature bound [Ma, 2023]

Let (M, d_g) be a complete Riemannian manifold with a lower Ricci curvature bound.
If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of a.c measures, then its barycenter $\overline{\mu}$ is a.c.

Sketch of proof, when $\mathbb{P} = \sum_{i=1}^{n} \lambda_i \delta_{\mu_i}$ and each μ_i has compact support

Similar to the case of displacement interpolation: locally Lipschitz + compactness

1. When μ_1 is a.c and μ_i's for $2 \leq i \leq n$ are Dirac measures, the optimal transport map from $\overline{\mu}$ to μ_1 is locally Lipschitz. `(See details later)`
2. Apply a divide-and-conquer `(conditional measure)` argument for the case when $\mu_i, 2 \leq i \leq n$ are discrete measures to retain the Lipschitz estimate.
3. Compactness and Rauch comparison theorem imply a uniform Lipschitz estimate for approximating sequences of general $\mu_i, i \leq 2 \leq n$.

How to prove absolute continuity

Absolute continuity and Ricci curvature bound [Ma, 2023]

Let (M, d_g) be a complete Riemannian manifold with a lower Ricci curvature bound.
If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of a.c measures, then its barycenter $\overline{\mu}$ is a.c.

Sketch of proof, when $\mathbb{P} = \sum_{i=1}^{n} \lambda_i \delta_{\mu_i}$ and each μ_i has compact support

Similar to the case of displacement interpolation: locally Lipschitz $+$ compactness

1. When μ_1 is a.c and μ_i's for $2 \leq i \leq n$ are Dirac measures, the optimal transport map from $\overline{\mu}$ to μ_1 is locally Lipschitz. `(See details later)`
2. Apply a divide-and-conquer `(conditional measure)` argument for the case when $\mu_i, 2 \leq i \leq n$ are discrete measures to retain the Lipschitz estimate.
3. Compactness and Rauch comparison theorem imply a uniform Lipschitz estimate for approximating sequences of general $\mu_i, i \leq 2 \leq n$.

How to prove absolute continuity

Absolute continuity and Ricci curvature bound [Ma, 2023]

Let (M, d_g) be a complete Riemannian manifold with a lower Ricci curvature bound.
If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of a.c measures, then its barycenter $\overline{\mu}$ is a.c.

Sketch of proof, when $\mathbb{P} = \sum_{i=1}^{n} \lambda_i \, \delta_{\mu_i}$ and each μ_i has compact support

Similar to the case of displacement interpolation: locally Lipschitz + compactness

1. When μ_1 is a.c and μ_i's for $2 \leq i \leq n$ are Dirac measures, the optimal transport map from $\overline{\mu}$ to μ_1 is locally Lipschitz. `(See details later)`
2. Apply a divide-and-conquer `(conditional measure)` argument for the case when $\mu_i, 2 \leq i \leq n$ are discrete measures to retain the Lipschitz estimate.
3. Compactness and Rauch comparison theorem imply a uniform Lipschitz estimate for approximating sequences of general $\mu_i, i \leq 2 \leq n$.

How to prove absolute continuity

Absolute continuity and Ricci curvature bound [Ma, 2023]

Let (M, d_g) be a complete Riemannian manifold with a lower Ricci curvature bound. If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of a.c measures, then its barycenter $\overline{\mu}$ is a.c.

Pass to the general case of $\mathbb{P}$ by consistency

Hessian equality for Wasserstein barycenters: let $\overline{\mu}$ be the unique a.c barycenter of $\sum_{i=1}^{n} \lambda_i \delta_{\mu_i}$ and let $\exp(-\nabla \phi_i)$ be the optimal transport map between $\overline{\mu}$ and μ_i, then

$$\sum_{i=1}^{n} \lambda_i \operatorname{Hess} \phi_i = 0.$$

How to prove absolute continuity

Absolute continuity and Ricci curvature bound [Ma, 2023]

Let (M, d_g) be a complete Riemannian manifold with a lower Ricci curvature bound. If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of a.c measures, then its barycenter $\overline{\mu}$ is a.c.

Pass to the general case of $\mathbb{P}$ by consistency

Hessian equality for Wasserstein barycenters: let $\overline{\mu}$ be the unique a.c barycenter of $\sum_{i=1}^{n} \lambda_i \delta_{\mu_i}$ and let $\exp(-\nabla \phi_i)$ be the optimal transport map between $\overline{\mu}$ and μ_i, then

$$\sum_{i=1}^{n} \lambda_i \operatorname{Hess} \phi_i \geq 0.$$

Approach of [Kim and Pass, 2017]: apply change of variable formula in the inequality and bound the density of $\overline{\mu}$ by a uniform upper bound of those of μ_i's.

How to prove absolute continuity

Absolute continuity and Ricci curvature bound [Ma, 2023]

Let (M, d_g) be a complete Riemannian manifold with a lower Ricci curvature bound.
If $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ gives mass to the set of a.c measures, then its barycenter $\overline{\mu}$ is a.c.

Pass to the general case of $\mathbb{P}$ by consistency

Hessian equality for Wasserstein barycenters: let $\overline{\mu}$ be the unique a.c barycenter of $\sum_{i=1}^{n} \lambda_i \delta_{\mu_i}$ and let $\exp(-\nabla \phi_i)$ be the optimal transport map between $\overline{\mu}$ and μ_i, then

$$\sum_{i=1}^{n} \lambda_i \operatorname{Hess} \phi_i = 0.$$

Our approach [Ma, 2023]: define nice functionals admitting finite values only for a.c measures, and bound them from above with the help of Souslin space theory.

Absolute continuity of Wasserstein barycenters of finitely many measures

Fix $\mathbb{P} = \sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$, where μ_1 is a.c with compact support and $\mu_i = \delta_{x_i}$ for $i \geq 2$. Its unique barycenter is $\overline{\mu} = B_{\#}\gamma$, where B is a measurable barycenter selection map and $\gamma = \mu_1 \otimes \delta_{x_2} \otimes \cdots \delta_{x_n}$ is the unique coupling of its marginals.

c-conjugating formulation of B

1. Define $c(x, y) := \frac{1}{2} d_g(x, y)^2$ and $h(y) := -\frac{1}{\lambda_1} \sum_{i=2}^n \lambda_i \, c(x_i, y)$
2. Given $x_1 \in M$, z is a barycenter of $\nu := \sum_{i=1}^n \lambda_i \, \delta_{x_i}$

$$\iff z \quad \text{reaches the infimum of} \quad 2\lambda_1 \inf_{y \in M} \{c(x_1, y) - h(y)\}$$

3. Define $X = \mathrm{supp}(\mu_1)$ and Y the set of barycenters of ν when x_1 runs through X. The map h is smooth on Y [Kim and Pass, 2015]. Set $F := \exp(-\nabla h)$.

$$z \in Y \text{ and } x_1 = F(z) \iff x_1 \in X \text{ and } z \text{ is a barycenter of } \nu$$

Conclusion: $F_{\#}\overline{\mu} = \mu_1$. Since F is Lipschitz, $\overline{\mu}$ is a.c.

Absolute continuity of Wasserstein barycenters of finitely many measures

Fix $\mathbb{P} = \sum_{i=1}^{n} \lambda_i \, \delta_{\mu_i}$, where μ_1 is a.c with compact support and $\mu_i = \delta_{x_i}$ for $i \geq 2$. Its unique barycenter is $\overline{\mu} = B_{\#}\gamma$, where B is a measurable barycenter selection map and $\gamma = \mu_1 \otimes \delta_{x_2} \otimes \cdots \delta_{x_n}$ is the unique coupling of its marginals.

c-conjugating formulation of B

1. Define $c(x,y) := \frac{1}{2} d_g(x,y)^2$ and $h(y) := -\frac{1}{\lambda_1} \sum_{i=2}^{n} \lambda_i \, c(x_i, y)$
2. Given $x_1 \in M$, z is a barycenter of $\nu := \sum_{i=1}^{n} \lambda_i \, \delta_{x_i}$

   $$\iff z \quad \text{reaches the infimum of} \quad 2\lambda_1 \inf_{y \in M} \{c(x_1, y) - h(y)\}$$

3. Define $X = \operatorname{supp}(\mu_1)$ and Y the set of barycenters of ν when x_1 runs through X. The map h is smooth on Y [Kim and Pass, 2015]. Set $F := \exp(-\nabla h)$.

   $$z \in Y \text{ and } x_1 = F(z) \iff x_1 \in X \text{ and } z \text{ is a barycenter of } \nu$$

Conclusion: $F_{\#}\overline{\mu} = \mu_1$. Since F is Lipschitz, $\overline{\mu}$ is a.c.

Absolute continuity of Wasserstein barycenters of finitely many measures

Fix $\mathbb{P} = \sum_{i=1}^{n} \lambda_i \, \delta_{\mu_i}$, where μ_1 is a.c with compact support and $\mu_i = \delta_{x_i}$ for $i \geq 2$.
Its unique barycenter is $\overline{\mu} = B_{\#}\gamma$, where B is a measurable barycenter selection map and $\gamma = \mu_1 \otimes \delta_{x_2} \otimes \cdots \delta_{x_n}$ is the unique coupling of its marginals.

c-conjugating formulation of B

1. Define $c(x, y) := \frac{1}{2} d_g(x, y)^2$ and $h(y) := -\frac{1}{\lambda_1} \sum_{i=2}^{n} \lambda_i \, c(x_i, y)$
2. Given $x_1 \in M$, z is a barycenter of $\nu := \sum_{i=1}^{n} \lambda_i \, \delta_{x_i}$

$$\iff z \quad \text{reaches the infimum of} \quad 2\lambda_1 \inf_{y \in M} \{c(x_1, y) - h(y)\}$$

3. Define $X = \operatorname{supp}(\mu_1)$ and Y the set of barycenters of ν when x_1 runs through X. The map h is smooth on Y [Kim and Pass, 2015]. Set $F := \exp(-\nabla h)$.

$$z \in Y \text{ and } x_1 = F(z) \iff x_1 \in X \text{ and } z \text{ is a barycenter of } \nu$$

Conclusion: $F_{\#}\overline{\mu} = \mu_1$. Since F is Lipschitz, $\overline{\mu}$ is a.c.

Absolute continuity of Wasserstein barycenters of finitely many measures

Fix $\mathbb{P} = \sum_{i=1}^{n} \lambda_i \, \delta_{\mu_i}$, where μ_1 is a.c with compact support and $\mu_i = \delta_{x_i}$ for $i \geq 2$.
Its unique barycenter is $\overline{\mu} = B_{\#}\gamma$, where B is a measurable barycenter selection map and $\gamma = \mu_1 \otimes \delta_{x_2} \otimes \cdots \delta_{x_n}$ is the unique coupling of its marginals.

c-conjugating formulation of B

1. Define $c(x, y) := \frac{1}{2} d_g(x, y)^2$ and $h(y) := -\frac{1}{\lambda_1} \sum_{i=2}^{n} \lambda_i \, c(x_i, y)$
2. Given $x_1 \in M$, z is a barycenter of $\nu := \sum_{i=1}^{n} \lambda_i \, \delta_{x_i}$ c-conjugation of h

$$\iff z \quad \text{reaches the infimum of} \quad 2\lambda_1 \inf_{y \in M} \{ c(x_1, y) - h(y) \}$$

3. Define $X = \operatorname{supp}(\mu_1)$ and Y the set of barycenters of ν when x_1 runs through X. The map h is smooth on Y [Kim and Pass, 2015]. Set $F := \exp(-\nabla h)$.

$$z \in Y \text{ and } x_1 = F(z) \iff x_1 \in X \text{ and } z \text{ is a barycenter of } \nu$$

Conclusion: $F_{\#}\overline{\mu} = \mu_1$. Since F is Lipschitz, $\overline{\mu}$ is a.c.

Absolute continuity of Wasserstein barycenters of finitely many measures

Fix $\mathbb{P} = \sum_{i=1}^{n} \lambda_i \, \delta_{\mu_i}$, where μ_1 is a.c with compact support and $\mu_i = \delta_{x_i}$ for $i \geq 2$. Its unique barycenter is $\overline{\mu} = B_{\#}\gamma$, where B is a measurable barycenter selection map and $\gamma = \mu_1 \otimes \delta_{x_2} \otimes \cdots \delta_{x_n}$ is the unique coupling of its marginals.

c-conjugating formulation of B

1. Define $c(x, y) := \frac{1}{2} d_g(x, y)^2$ and $h(y) := -\frac{1}{\lambda_1} \sum_{i=2}^{n} \lambda_i \, c(x_i, y)$
2. Given $x_1 \in M$, z is a barycenter of $\nu := \sum_{i=1}^{n} \lambda_i \, \delta_{x_i}$

$$\iff z \quad \text{reaches the infimum of} \quad 2\lambda_1 \inf_{y \in M} \{c(x_1, y) - h(y)\}$$

3. Define $X = \mathrm{supp}(\mu_1)$ and Y the set of barycenters of ν when x_1 runs through X. The map h is smooth on Y [Kim and Pass, 2015]. Set $F := \exp(-\nabla h)$.

$$z \in Y \text{ and } x_1 = F(z) \iff x_1 \in X \text{ and } z \text{ is a barycenter of } \nu$$

Conclusion: $F_{\#}\overline{\mu} = \mu_1$. Since F is Lipschitz, $\overline{\mu}$ is a.c.

Absolute continuity of Wasserstein barycenters of finitely many measures

Fix $\mathbb{P} = \sum_{i=1}^n \lambda_i \, \delta_{\mu_i}$, where μ_1 is a.c with compact support and $\mu_i = \delta_{x_i}$ for $i \geq 2$.
Its unique barycenter is $\overline{\mu} = B_{\#}\gamma$, where B is a measurable barycenter selection map and $\gamma = \mu_1 \otimes \delta_{x_2} \otimes \cdots \delta_{x_n}$ is the unique coupling of its marginals.

c-conjugating formulation of B

1. Define $c(x,y) := \frac{1}{2} d_g(x,y)^2$ and $h(y) := -\frac{1}{\lambda_1} \sum_{i=2}^n \lambda_i \, c(x_i, y)$
2. Given $x_1 \in M$, z is a barycenter of $\nu := \sum_{i=1}^n \lambda_i \, \delta_{x_i}$

   $$\iff z \quad \text{reaches the infimum of} \quad 2\lambda_1 \inf_{y \in M} \{ c(x_1, y) - h(y) \}$$
3. Define $X = \mathrm{supp}(\mu_1)$ and Y the set of barycenters of ν when x_1 runs through X. The map h is smooth on Y [Kim and Pass, 2015]. Set $F := \exp(-\nabla h)$.

$$z \in Y \text{ and } x_1 = F(z) \iff x_1 \in X \text{ and } z \text{ is a barycenter of } \nu$$

Conclusion: $F_{\#}\overline{\mu} = \mu_1$. Since F is Lipschitz, $\overline{\mu}$ is a.c.

Displacement functionals for Wasserstein barycenters

Assumptions and notation for the functional $\mathcal{G} : f \cdot \mathrm{Vol} \mapsto \int_M G(f)\,\mathrm{d}\,\mathrm{Vol}$

1. $m = \dim(M)$, $\mathrm{Ric}_M \geq -(m-1)K\, g_M$; $\mathbb{P} = \sum_{i=1}^n \lambda_i\, \delta_{\mu_i}$, μ_i has compact support.
2. $\mu_i = g_i\ \mathrm{Vol}, 1 \leq i \leq k$ are a.c; the unique barycenter $\overline{\mu} = f\,\mathrm{Vol}$ of $\mathbb{P}$ is a.c.
3. $G : \mathbb{R}^+ \to \mathbb{R}$ with $G(0) = 0$ such that $H(x) := G(e^x)e^{-x}$ is $\mathcal{C}^1$ with non-negative derivatives bounded above by $L_H > 0$.

Define $\Lambda := \sum_{i=1}^k \lambda_i$, then

$$\mathcal{G}(\overline{\mu}) := \int_M G(f)\,\mathrm{d}\,\mathrm{Vol} \leq \sum_{i=1}^k \frac{\lambda_i}{\Lambda} \int_M G(g_i)\,\mathrm{d}\,\mathrm{Vol} + \frac{L_H K}{2\Lambda} \mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)\,.$$

Displacement functionals for Wasserstein barycenters

Assumptions and notation for the functional $\mathcal{G} : f \cdot \mathrm{Vol} \mapsto \int_M G(f)\,\mathrm{d}\,\mathrm{Vol}$

1. $m := \dim(M)$, $\mathrm{Ric}_M \geq -(m-1)K\, g_M$; $\mathbb{P} := \sum_{i=1}^{n} \lambda_i\, \delta_{\mu_i}$, μ_i has compact support.
2. $\mu_i = g_i\, \mathrm{Vol}, 1 \leq i \leq k$ are a.c; the unique barycenter $\overline{\mu} = f\, \mathrm{Vol}$ of $\mathbb{P}$ is a.c.
3. $G : \mathbb{R}^+ \to \mathbb{R}$ with $G(0) = 0$ such that $H(x) := G(e^x)e^{-x}$ is $\mathcal{C}^1$ with non-negative derivatives bounded above by $L_H > 0$.

Define $\Lambda := \sum_{i=1}^{k} \lambda_i$, then

$$\mathcal{G}(\overline{\mu}) := \int_M G(f)\,\mathrm{d}\,\mathrm{Vol} \leq \sum_{i=1}^{k} \frac{\lambda_i}{\Lambda} \int_M G(g_i)\,\mathrm{d}\,\mathrm{Vol} + \frac{L_H K}{2\Lambda} \mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)\,.$$

Displacement functionals for Wasserstein barycenters

Assumptions and notation for the functional $\mathcal{G} : f \cdot \mathrm{Vol} \mapsto \int_M G(f)\, \mathrm{d}\,\mathrm{Vol}$

1. $m := \dim(M)$, $\mathrm{Ric}_M \geq -(m-1)K\, g_M$; $\mathbb{P} := \sum_{i=1}^n \lambda_i\, \delta_{\mu_i}$, μ_i has compact support.
2. $\mu_i = g_i\, \mathrm{Vol}, 1 \leq i \leq k$ are a.c; the unique barycenter $\overline{\mu} = f\, \mathrm{Vol}$ of $\mathbb{P}$ is a.c.
3. $G : \mathbb{R}^+ \to \mathbb{R}$ with $G(0) = 0$ such that $H(x) := G(e^x)e^{-x}$ is $\mathcal{C}^1$ with non-negative derivatives bounded above by $L_H > 0$.

Define $\Lambda := \sum_{i=1}^k \lambda_i$, then

$$\mathcal{G}(\overline{\mu}) \leq \sum_{i=1}^k \frac{\lambda_i}{\Lambda} \mathcal{G}(\mu_i) + \frac{L_H K}{2\Lambda} \mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)\, .$$

Special case: curvature-dimension condition

Take $G(x) := x \log x$, $n = k = 2$, $\Lambda = L_H = 1$. Set $\lambda = \lambda_1$ and $\mathrm{Ent} = \mathcal{G}$, then

$$\mathrm{Ent}(\overline{\mu}) \leq \lambda\, \mathrm{Ent}(\mu_1) + (1-\lambda)\, \mathrm{Ent}(\mu_2) + \frac{K}{2} \lambda(1-\lambda)\, W(\mu_1, \mu_2)^2 + \frac{m^2}{2} + m\, .$$

Displacement functionals for Wasserstein barycenters

Assumptions and notation for the functional $\mathcal{G} : f \cdot \mathrm{Vol} \mapsto \int_M G(f)\, \mathrm{d}\,\mathrm{Vol}$

1. $m := \dim(M)$, $\mathrm{Ric}_M \geq -(m-1)K\, g_M$; $\mathbb{P} := \sum_{i=1}^n \lambda_i\, \delta_{\mu_i}$, μ_i has compact support.
2. $\mu_i = g_i\, \mathrm{Vol}, 1 \leq i \leq k$ are a.c; the unique barycenter $\overline{\mu} = f\, \mathrm{Vol}$ of $\mathbb{P}$ is a.c.
3. $G : \mathbb{R}^+ \to \mathbb{R}$ with $G(0) = 0$ such that $H(x) := G(e^x)e^{-x}$ is $\mathcal{C}^1$ with non-negative derivatives bounded above by $L_H > 0$.

Define $\Lambda := \sum_{i=1}^k \lambda_i$, then

$$\mathcal{G}(\overline{\mu}) \leq \sum_{i=1}^k \frac{\lambda_i}{\Lambda}\mathcal{G}(\mu_i) + \frac{L_H K}{2\Lambda}\mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)\,.$$

Difference from classical displacement functionals

Gradient flow theory `(first-order)` and displacement convexity `(second-order)` gives that

$$\mathcal{G}(\mu_i) \geq \mathcal{G}(\overline{\mu}) + \int_M \Delta\phi_i\, H'(\log f)\, \mathrm{d}\,\overline{\mu} - \frac{L_H\, K}{2} W_2(\overline{\mu}, \mu_i)^2, \quad 1 \leq i \leq k.$$

Preservation of absolute continuity when passing to the limit

Reminder of the problem setting

We approximate a general measure $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ with $\mathbb{P}_j$. After proving that the barycenter $\overline{\mu}_j$ of $\mathbb{P}_j$ is a.c, how to show that the barycenter $\overline{\mu} = \lim \overline{\mu}_j$ of $\mathbb{P}$ is also a.c?

Use displacement functionals $\mathcal{G}$ admitting finite values only for a.c measures

1. Assume G is in addition super-linear and convex, then $\mathcal{G}$ is lower semi-continuous;
2. Bound $\{\mathcal{G}(\overline{\mu}_j)\}_{j \geq 1}$ from above, for which we use the displacement inequality;
3. By choosing the sequence $\mathbb{P}_j$ properly, it reduces to show that $\mathbb{P}$ gives mass to a $\mathrm{B}(G, L)$ set, the set of a.c measures whose values under $\mathcal{G}$ are bounded by $L > 0$;
4. Compact sets w.r.t. the $\sigma(L^1, L^\infty)$ topology are $B(G, L)$ sets;
5. Souslin space theory implies that several different topologies of a.c measures generate the same Borel sets, on which $\mathbb{P}$ is a Radon measure.

Preservation of absolute continuity when passing to the limit

Reminder of the problem setting

We approximate a general measure $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ with $\mathbb{P}_j$. After proving that the barycenter $\overline{\mu}_j$ of $\mathbb{P}_j$ is a.c, how to show that the barycenter $\overline{\mu} = \lim \overline{\mu}_j$ of $\mathbb{P}$ is also a.c?

Use displacement functionals $\mathcal{G}$ admitting finite values only for a.c measures

1. Assume G is in addition super-linear and convex, then $\mathcal{G}$ is lower semi-continuous;
2. Bound $\{\mathcal{G}(\overline{\mu}_j)\}_{j\geq 1}$ from above, for which we use the displacement inequality;
3. By choosing the sequence $\mathbb{P}_j$ properly, it reduces to show that $\mathbb{P}$ gives mass to a $\mathrm{B}(G, L)$ set, the set of a.c measures whose values under $\mathcal{G}$ are bounded by $L > 0$;
4. Compact sets w.r.t. the $\sigma(L^1, L^\infty)$ topology are $B(G, L)$ sets;
5. Souslin space theory implies that several different topologies of a.c measures generate the same Borel sets, on which $\mathbb{P}$ is a Radon measure.

Preservation of absolute continuity when passing to the limit

Reminder of the problem setting

We approximate a general measure $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ with $\mathbb{P}_j$. After proving that the barycenter $\overline{\mu}_j$ of $\mathbb{P}_j$ is a.c, how to show that the barycenter $\overline{\mu} = \lim \overline{\mu}_j$ of $\mathbb{P}$ is also a.c?

Use displacement functionals $\mathcal{G}$ admitting finite values only for a.c measures

1. Assume G is in addition super-linear and convex, then $\mathcal{G}$ is lower semi-continuous;
2. Bound $\{\mathcal{G}(\overline{\mu}_j)\}_{j \geq 1}$ from above, for which we use the displacement inequality;
3. By choosing the sequence $\mathbb{P}_j$ properly, it reduces to show that $\mathbb{P}$ gives mass to a $\mathrm{B}(G, L)$ set, the set of a.c measures whose values under $\mathcal{G}$ are bounded by $L > 0$;
4. Compact sets w.r.t. the $\sigma(L^1, L^\infty)$ topology are $B(G, L)$ sets;
5. Souslin space theory implies that several different topologies of a.c measures generate the same Borel sets, on which $\mathbb{P}$ is a Radon measure.

Preservation of absolute continuity when passing to the limit

Reminder of the problem setting

We approximate a general measure $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ with $\mathbb{P}_j$. After proving that the barycenter $\overline{\mu}_j$ of $\mathbb{P}_j$ is a.c, how to show that the barycenter $\overline{\mu} = \lim \overline{\mu}_j$ of $\mathbb{P}$ is also a.c?

Use displacement functionals $\mathcal{G}$ admitting finite values only for a.c measures

1. Assume G is in addition super-linear and convex, then $\mathcal{G}$ is lower semi-continuous;
2. Bound $\{\mathcal{G}(\overline{\mu}_j)\}_{j \geq 1}$ from above, for which we use the displacement inequality;
3. By choosing the sequence $\mathbb{P}_j$ properly, it reduces to show that $\mathbb{P}$ gives mass to a $\mathrm{B}(G, L)$ set, the set of a.c measures whose values under $\mathcal{G}$ are bounded by $L > 0$;
4. Compact sets w.r.t. the $\sigma(L^1, L^\infty)$ topology are $B(G, L)$ sets;
5. Souslin space theory implies that several different topologies of a.c measures generate the same Borel sets, on which $\mathbb{P}$ is a Radon measure.

Preservation of absolute continuity when passing to the limit

Reminder of the problem setting

We approximate a general measure $\mathbb{P} \in \mathcal{W}(\mathcal{W}(M))$ with $\mathbb{P}_j$. After proving that the barycenter $\overline{\mu}_j$ of $\mathbb{P}_j$ is a.c, how to show that the barycenter $\overline{\mu} = \lim \overline{\mu}_j$ of $\mathbb{P}$ is also a.c?

Use displacement functionals $\mathcal{G}$ admitting finite values only for a.c measures

1. Assume G is in addition super-linear and convex, then $\mathcal{G}$ is lower semi-continuous;
2. Bound $\{\mathcal{G}(\overline{\mu}_j)\}_{j\geq 1}$ from above, for which we use the displacement inequality;
3. By choosing the sequence $\mathbb{P}_j$ properly, it reduces to show that $\mathbb{P}$ gives mass to a $\mathrm{B}(G, L)$ set, the set of a.c measures whose values under $\mathcal{G}$ are bounded by $L > 0$;
4. Compact sets w.r.t. the $\sigma(L^1, L^\infty)$ topology are $B(G, L)$ sets;
5. Souslin space theory implies that several different topologies of a.c measures generate the same Borel sets, on which $\mathbb{P}$ is a Radon measure.

References

Agueh, M. and Carlier, G. (2011). Barycenters in the Wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924.

Cordero-Erausquin, D., McCann, R. J., and Schmuckenschläger, M. (2001). A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Inventiones mathematicae*, 146(2):219–257.

Kim, Y.-H. and Pass, B. (2015). Multi-marginal optimal transport on Riemannian manifolds. *American Journal of Mathematics*, 137(4):1045–1060.

Kim, Y.-H. and Pass, B. (2017). Wasserstein barycenters over Riemannian manifolds. *Advances in Mathematics*, 307:640–683.

Le Gouic, T. and Loubes, J.-M. (2017). Existence and consistency of Wasserstein barycenters. *Probability Theory and Related Fields*, 168(3):901–917.

Ma, J. (2023). Absolute continuity of Wasserstein barycenters on manifolds with a lower Ricci curvature bound. *arXiv preprint arXiv:2310.13832*.

Justifications for the generalized displacement functionals

$$\mathcal{G}(\overline{\mu}) \leq \sum_{i=1}^{k} \frac{\lambda_i}{\Lambda}\mathcal{G}(\mu_i) + \frac{L_H K}{2\Lambda}\mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)$$

Step 1, change of variables

Denote by F_i the optimal transport map from $\overline{\mu}$ to μ_i, by $\mathrm{Jac}\, F_i$ the Jacobian of F_i. Since $f = g(F_i)\,\mathrm{Jac}\, F_i$, $\quad \mathcal{G}(\mu_i) = \int_M H(\log f + l_i)\,\mathrm{d}\,\overline{\mu}$, where $l_i := -\log \mathrm{Jac}\, F_i$.

Justifications for the generalized displacement functionals

$$\mathcal{G}(\overline{\mu}) \leq \sum_{i=1}^{k} \frac{\lambda_i}{\Lambda} \mathcal{G}(\mu_i) + \frac{L_H K}{2\Lambda} \mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)$$

Step 1, change of variables

Denote by F_i the optimal transport map from $\overline{\mu}$ to μ_i, by $\mathrm{Jac}\, F_i$ the Jacobian of F_i.
Since $f = g(F_i)\,\mathrm{Jac}\, F_i$, $\quad \mathcal{G}(\mu_i) = \int_M H(\log f + l_i)\,\mathrm{d}\,\overline{\mu}$, where $l_i := -\log \mathrm{Jac}\, F_i$.
By McCann-Brenier theorem, $F_i = \exp(-\nabla \phi_i)$ with ϕ_i a c-concave function.

Justifications for the generalized displacement functionals

$$\mathcal{G}(\overline{\mu}) \leq \sum_{i=1}^{k} \frac{\lambda_i}{\Lambda}\mathcal{G}(\mu_i) + \frac{L_H K}{2\Lambda}\mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)$$

Step 1, change of variables

Denote by F_i the optimal transport map from $\overline{\mu}$ to μ_i, by $\mathrm{Jac}\, F_i$ the Jacobian of F_i.
Since $f = g(F_i)\,\mathrm{Jac}\, F_i$, $\quad \mathcal{G}(\mu_i) = \int_M H(\log f + l_i)\,\mathrm{d}\,\overline{\mu}$, where $l_i := -\log \mathrm{Jac}\, F_i$.
By McCann-Brenier theorem, $F_i = \exp(-\nabla\phi_i)$ with ϕ_i a c-concave function.

Step 2, apply Ricci curvature bound

Jacobi equation for $\mathrm{d}\exp(-\nabla\phi_i)$ implies $l_i \geq \Delta\phi_i - K\|\nabla\phi_i\|^2/2$ for $1 \leq i \leq k$.
Second variation formula implies $m + m^2/2 \geq \Delta\phi_i - K\|\nabla\phi_i\|^2/2$. [Cordero-Erausquin et al., 2001]

Justifications for the generalized displacement functionals

$$\mathcal{G}(\overline{\mu}) \leq \sum_{i=1}^{k} \frac{\lambda_i}{\Lambda}\mathcal{G}(\mu_i) + \frac{L_H K}{2\Lambda}\mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)$$

Step 1, change of variables

Denote by F_i the optimal transport map from $\overline{\mu}$ to μ_i, by $\operatorname{Jac} F_i$ the Jacobian of F_i.
Since $f = g(F_i)\operatorname{Jac} F_i$, $\quad \mathcal{G}(\mu_i) = \int_M H(\log f + l_i)\,\mathrm{d}\,\overline{\mu}$, where $l_i := -\log \operatorname{Jac} F_i$.
By McCann-Brenier theorem, $F_i = \exp(-\nabla\phi_i)$ with ϕ_i a c-concave function.

Step 2, apply Ricci curvature bound

Jacobi equation for $\mathrm{d}\exp(-\nabla\phi_i)$ implies $l_i \geq \Delta\phi_i - K\|\nabla\phi_i\|^2/2$ for $1 \leq i \leq k$.
Second variation formula implies $m + m^2/2 \geq \Delta\phi_i - K\|\nabla\phi_i\|^2/2$. [Cordero-Erausquin et al., 2001]

Step 3, apply assumptions on H

$H(\log f + l_i) - H(\log f) \geq L_H(\Delta\phi_i - K\|\nabla\phi_i\|^2/2) - L_H(m + m^2/2)$.

Justifications for the generalized displacement functionals

$$\mathcal{G}(\overline{\mu}) \leq \sum_{i=1}^{k} \frac{\lambda_i}{\Lambda} \mathcal{G}(\mu_i) + \frac{L_H K}{2\Lambda} \mathbb{W}(\mathbb{P}, \delta_{\overline{\mu}})^2 + \frac{L_H}{2\Lambda}(m^2 + 2m)$$

Step 1, change of variables

Denote by F_i the optimal transport map from $\overline{\mu}$ to μ_i, by $\operatorname{Jac} F_i$ the Jacobian of F_i.
Since $f = g(F_i) \operatorname{Jac} F_i$, $\quad \mathcal{G}(\mu_i) = \int_M H(\log f + l_i) \, \mathrm{d}\, \overline{\mu}$, where $l_i := -\log \operatorname{Jac} F_i$.
By McCann-Brenier theorem, $F_i = \exp(-\nabla \phi_i)$ with ϕ_i a c-concave function.

Step 2, apply Ricci curvature bound

Jacobi equation for $\mathrm{d} \exp(-\nabla \phi_i)$ implies $l_i \geq \Delta \phi_i - K \|\nabla \phi_i\|^2/2$ for $1 \leq i \leq k$.
Second variation formula implies $m + m^2/2 \geq \Delta \phi_i - K \|\nabla \phi_i\|^2/2$. [Cordero-Erausquin et al., 2001]

Step 3, apply assumptions on H

$H(\log f + l_i) - H(\log f) \geq L_H(\Delta \phi_i - K \|\nabla \phi_i\|^2/2) - L_H(m + m^2/2)$.

Step 4, integrate and apply the Hessian equality

The Hessian equality $\sum_{i=1}^{n} \lambda_i \operatorname{Hess}_x \phi_i = 0$ implies $\sum_{i=1}^{n} \lambda_i \Delta \phi_i(x) = 0$.