Quantitative Stability of the Pushforward Operation by an Optimal Transport Map

Alex Delalande

Joint work with Guillaume Carlier and Quentin Mérigot

January 2024

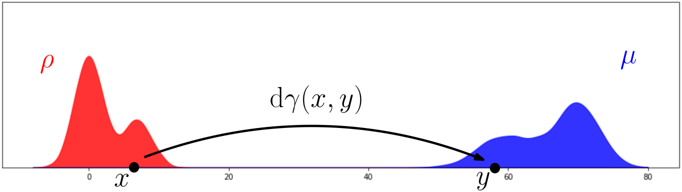
Introduction

Quadratic Optimal Transport problem (Monge, 1781; Kantorovich, 1942):

- Given $\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d)$, solve

$$\inf_{\gamma \in \Gamma(\rho,\mu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^2 \, \mathrm{d}\gamma(x,y),$$

where $\Gamma(\rho, \mu) = \{\gamma \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) \mid \forall A \subset \Omega, \gamma(A \times \mathbb{R}^d) = \rho(A), \gamma(\mathbb{R}^d \times A) = \mu(A)\}$.



- Optimal γ always **exists** and is **unique if ρ is absolutely continuous**.

Theorem (Brenier, 1987): If ρ is **absolutely continuous**, solution is induced by a **unique map** $T : \mathbb{R}^d \to \mathbb{R}^d$ satisfying $T_{\#}\rho = \mu$, **characterized by** $T = \nabla\phi$ **with** ϕ **convex**.

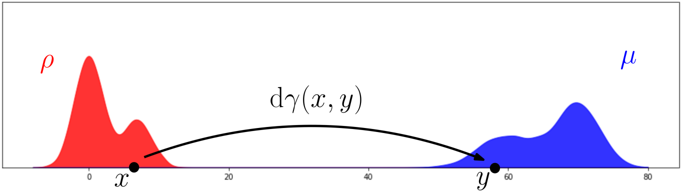
Introduction

Quadratic Optimal Transport problem (Monge, 1781; Kantorovich, 1942):

- Given $\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d)$, solve

$$\inf_{\gamma \in \Gamma(\rho,\mu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^2 \, \mathrm{d}\gamma(x, y),$$

where $\Gamma(\rho, \mu) = \{\gamma \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) \mid \forall A \subset \Omega, \gamma(A \times \mathbb{R}^d) = \rho(A), \gamma(\mathbb{R}^d \times A) = \mu(A)\}$.



- Optimal γ always **exists** and is **unique if ρ is absolutely continuous**.

Theorem (Brenier, 1987): If ρ is **absolutely continuous**, solution is induced by a **unique map** $T : \mathbb{R}^d \to \mathbb{R}^d$ satisfying $T_{\#}\rho = \mu$, **characterized by** $T = \nabla\phi$ **with** ϕ **convex**.

Introduction

- $2-$***Wasserstein distance*** between ρ and μ:

$$\boxed{\mathrm{W}_2(\rho,\mu) := \left(\inf_{\gamma\in\Gamma(\rho,\mu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2 \,\mathrm{d}\gamma(x,y)\right)^{1/2}.}$$

- $2-$***Wasserstein space***: $(\mathcal{P}_2(\mathbb{R}^d), \mathrm{W}_2)$.

$\rightarrow$ Geodesic distance, interpolations, barycenters, gradient flows...

$\rightarrow$ Riemannian interpretation of $(\mathcal{P}_2(\mathbb{R}^d), \mathrm{W}_2)$ (Otto, 2001; Ambrosio, Gigli, Savaré, 2004).

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_{\#}\rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho\mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \text{ convex}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi\|\rho\rangle + \langle\phi^*\|\mu\rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_{\#}\rho \in \mathcal{P}_2(\mathbb{R}^d)$

$\to$ *Tangent vectors* are *directed* by gradients of convex functions.

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_{\#}\rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho\mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \text{ convex}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi\|\rho\rangle + \langle\phi^*\|\mu\rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_{\#}\rho \in \mathcal{P}_2(\mathbb{R}^d)$

$\rightarrow$ *Tangent vectors* are *directed* by gradients of convex functions.

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_{\#}\rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho\mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \text{ convex}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi\|\rho\rangle + \langle\phi^*\|\mu\rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_{\#}\rho \in \mathcal{P}_2(\mathbb{R}^d)$

$\rightarrow$ *Tangent vectors* are *directed* by gradients of convex functions.

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_{\#}\rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho\mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \text{ convex}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi\|\rho\rangle + \langle\phi^*\|\mu\rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_{\#}\rho \in \mathcal{P}_2(\mathbb{R}^d)$

→ *Tangent vectors* are *directed* by gradients of convex functions.

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_\#\rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho\mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \text{ **convex**}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi\|\rho\rangle + \langle\phi^*\|\mu\rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_\#\rho \in \mathcal{P}_2(\mathbb{R}^d)$

$\to$ *Tangent vectors* are *directed* by gradients of convex functions.

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_\#\rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho\mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \textbf{ convex}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi\|\rho\rangle + \langle\phi^*\|\mu\rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_\#\rho \in \mathcal{P}_2(\mathbb{R}^d)$

→ *Tangent vectors* are *directed* by gradients of convex functions.

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_\# \rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho \mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \textbf{ convex}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho \mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \textbf{ convex}} \langle \phi \vert \rho \rangle + \langle \phi^* \vert \mu \rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_\#\rho \in \mathcal{P}_2(\mathbb{R}^d)$

$\rightarrow$ *Tangent vectors* are *directed* by gradients of convex functions.

Introduction

- Riemannian interpretation of $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$:

$$\mathrm{W}_2(\rho,\mu) = \|\nabla\phi - \mathrm{id}\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \quad \forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex and s.t. } (\nabla\phi)_{\#}\rho = \mu.$$

Point	$\mu \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$
Geodesic distance	$\mathrm{W}_2(\mu,\nu)$
Tangent space	$\mathcal{T}_\rho \mathcal{P}_2^{a.c.}(\mathbb{R}^d) = \overline{\{\lambda(\nabla\phi - \mathrm{id}) \mid \lambda > 0, \phi \textbf{ convex}\}}^{\mathrm{L}^2(\rho,\mathbb{R}^d)}$
Inverse exponential map	$\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id \in \mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$, where $\phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi\|\rho\rangle + \langle\phi^*\|\mu\rangle$
Exponential map	$\exp_\rho(\nabla\phi - \mathrm{id}) = (\nabla\phi)_{\#}\rho \in \mathcal{P}_2(\mathbb{R}^d)$

$\rightarrow$ *Tangent vectors* are *directed* by gradients of convex functions.

Problem statement

Let $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be a ***fixed***, proper and continuous **convex** function.

On $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$, what is the regularity of the map

$$\rho \mapsto (\nabla\phi)_{\#}\rho \quad \textbf{?}$$

I.e., can we have bounds of the type

$$\forall \rho, \tilde{\rho} \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d), \quad \mathrm{W}_2((\nabla\phi)_{\#}\rho, (\nabla\phi)_{\#}\tilde{\rho}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^{\alpha} \quad \textbf{?}$$

More generally, let $\rho, \tilde{\rho} \in \mathcal{P}_2(\mathbb{R}^d)$ and $\gamma, \tilde{\gamma} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ such that

$$\begin{cases} (p_1)_{\#}\gamma = \rho, \\ \mathrm{spt}(\gamma) \subset \partial\phi, \end{cases} \qquad \begin{cases} (p_1)_{\#}\tilde{\gamma} = \tilde{\rho}, \\ \mathrm{spt}(\tilde{\gamma}) \subset \partial\phi. \end{cases}$$

$$\to \mathrm{W}_2((p_2)_{\#}\gamma, (p_2)_{\#}\tilde{\gamma}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^{\alpha} \quad \textbf{?}$$

Problem statement

Let $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be a ***fixed***, proper and continuous **convex** function.

On $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$, what is the regularity of the map

$$\rho \mapsto (\nabla\phi)_{\#}\rho \quad \textbf{?}$$

I.e., can we have bounds of the type

$$\forall \rho, \tilde{\rho} \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d), \quad \mathrm{W}_2((\nabla\phi)_{\#}\rho, (\nabla\phi)_{\#}\tilde{\rho}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^{\alpha} \quad \textbf{?}$$

More generally, let $\rho, \tilde{\rho} \in \mathcal{P}_2(\mathbb{R}^d)$ and $\gamma, \tilde{\gamma} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ such that

$$\begin{cases} (p_1)_{\#}\gamma = \rho, \\ \mathrm{spt}(\gamma) \subset \partial\phi, \end{cases} \qquad \begin{cases} (p_1)_{\#}\tilde{\gamma} = \tilde{\rho}, \\ \mathrm{spt}(\tilde{\gamma}) \subset \partial\phi. \end{cases}$$

$$\to \mathrm{W}_2((p_2)_{\#}\gamma, (p_2)_{\#}\tilde{\gamma}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^{\alpha} \quad ?$$

Problem statement

Let $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be a ***fixed***, proper and continuous **convex** function.

On $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$, what is the regularity of the map

$$\rho \mapsto (\nabla\phi)_\# \rho \quad \textbf{?}$$

I.e., can we have bounds of the type

$$\forall \rho, \tilde{\rho} \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d), \quad \mathrm{W}_2((\nabla\phi)_\#\rho, (\nabla\phi)_\#\tilde{\rho}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^\alpha \quad \textbf{?}$$

More generally, let $\rho, \tilde{\rho} \in \mathcal{P}_2(\mathbb{R}^d)$ and $\gamma, \tilde{\gamma} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ such that

$$\begin{cases} (p_1)_\#\gamma = \rho, \\ \mathrm{spt}(\gamma) \subset \partial\phi, \end{cases} \qquad \begin{cases} (p_1)_\#\tilde{\gamma} = \tilde{\rho}, \\ \mathrm{spt}(\tilde{\gamma}) \subset \partial\phi. \end{cases}$$

$$\to \mathrm{W}_2((p_2)_\#\gamma, (p_2)_\#\tilde{\gamma}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^\alpha \quad \textbf{?}$$

Problem statement

Let $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be a ***fixed***, proper and continuous **convex** function.

On $(\mathcal{P}_2^{a.c.}(\mathbb{R}^d), \mathrm{W}_2)$**, what is the regularity of the map**

$$\rho \mapsto (\nabla\phi)_{\#}\rho \quad \textbf{?}$$

I.e., can we have bounds of the type

$$\forall \rho, \tilde{\rho} \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d), \quad \mathrm{W}_2((\nabla\phi)_{\#}\rho, (\nabla\phi)_{\#}\tilde{\rho}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^{\alpha} \quad \textbf{?}$$

More generally, let $\rho, \tilde{\rho} \in \mathcal{P}_2(\mathbb{R}^d)$ and $\gamma, \tilde{\gamma} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ such that

$$\begin{cases} (p_1)_{\#}\gamma = \rho, \\ \mathrm{spt}(\gamma) \subset \partial\phi, \end{cases} \qquad \begin{cases} (p_1)_{\#}\tilde{\gamma} = \tilde{\rho}, \\ \mathrm{spt}(\tilde{\gamma}) \subset \partial\phi. \end{cases}$$

$$\to \mathrm{W}_2((p_2)_{\#}\gamma, (p_2)_{\#}\tilde{\gamma}) \leq C\mathrm{W}_2(\rho, \tilde{\rho})^{\alpha} \quad \textbf{?}$$

Motivations

1. Numerical resolution of optimal transport

- By Kantorovich's duality, with $\psi^* = \sup_y \langle \cdot | y \rangle - \psi(y)$:

$$\min_{\gamma \in \Gamma(\rho,\mu)} \int \|x-y\|^2 \, \mathrm{d}\gamma(x,y) \equiv \min_{\psi:\mathbb{R}^d \to \mathbb{R}} \underbrace{\int \psi^* \mathrm{d}\rho}_{:=K(\psi)} + \int \psi \mathrm{d}\mu.$$

- Gradient of K:

$$\nabla K(\psi) = -(\nabla \psi^*)_{\#}\rho.$$

- Approximation of $\nabla K(\psi)$:

$$\tilde{\rho} := \frac{1}{N}\sum_i \delta_{x_i} \quad \to \quad \widetilde{\nabla K}(\psi) := -\frac{1}{N}\sum_i \delta_{y_i}, \text{ where } y_i \in \partial \psi^*(x_i).$$

- Quality of approximation?

$$\mathrm{W}_2(\nabla K(\psi), \widetilde{\nabla K}(\psi)) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?$$

Motivations

1. Numerical resolution of optimal transport

- By Kantorovich's duality, with $\psi^* = \sup_y \langle \cdot | y \rangle - \psi(y)$:

$$\min_{\gamma \in \Gamma(\rho,\mu)} \int \|x-y\|^2 \,\mathrm{d}\gamma(x,y) \equiv \min_{\psi:\mathbb{R}^d \to \mathbb{R}} \underbrace{\int \psi^* \mathrm{d}\rho + \int \psi \mathrm{d}\mu}_{:=K(\psi)}.$$

- Gradient of K:

$$\nabla K(\psi) = -(\nabla \psi^*)_{\#}\rho.$$

- Approximation of $\nabla K(\psi)$:

$$\tilde{\rho} := \frac{1}{N}\sum_i \delta_{x_i} \quad \to \quad \widetilde{\nabla K}(\psi) := -\frac{1}{N}\sum_i \delta_{y_i}, \text{ where } y_i \in \partial\psi^*(x_i).$$

- Quality of approximation?

$$\mathrm{W}_2(\nabla K(\psi), \widetilde{\nabla K}(\psi)) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?$$

Motivations

1. Numerical resolution of optimal transport

- By Kantorovich's duality, with $\psi^* = \sup_y \langle \cdot | y \rangle - \psi(y)$:

$$\min_{\gamma \in \Gamma(\rho,\mu)} \int \|x-y\|^2 \mathrm{d}\gamma(x,y) \equiv \min_{\psi:\mathbb{R}^d \to \mathbb{R}} \underbrace{\int \psi^* \mathrm{d}\rho}_{:=K(\psi)} + \int \psi \mathrm{d}\mu.$$

- Gradient of K:

$$\nabla K(\psi) = -(\nabla \psi^*)_{\#}\rho.$$

- Approximation of $\nabla K(\psi)$:

$$\tilde{\rho} := \frac{1}{N}\sum_i \delta_{x_i} \quad \to \quad \widetilde{\nabla K}(\psi) := -\frac{1}{N}\sum_i \delta_{y_i}, \text{ where } y_i \in \partial\psi^*(x_i).$$

- Quality of approximation?

$$\mathrm{W}_2(\nabla K(\psi), \widetilde{\nabla K}(\psi)) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?$$

Motivations

1. Numerical resolution of optimal transport

- By Kantorovich's duality, with $\psi^* = \sup_y \langle \cdot | y \rangle - \psi(y)$:

$$\min_{\gamma \in \Gamma(\rho,\mu)} \int \|x-y\|^2 \,\mathrm{d}\gamma(x,y) \equiv \min_{\psi:\mathbb{R}^d \to \mathbb{R}} \underbrace{\int \psi^* \mathrm{d}\rho}_{:=K(\psi)} + \int \psi \mathrm{d}\mu.$$

- Gradient of K:

$$\nabla K(\psi) = -(\nabla \psi^*)_{\#}\rho.$$

- Approximation of $\nabla K(\psi)$:

$$\tilde{\rho} := \frac{1}{N}\sum_i \delta_{x_i} \quad \to \quad \widetilde{\nabla K}(\psi) := -\frac{1}{N}\sum_i \delta_{y_i}, \text{ where } y_i \in \partial \psi^*(x_i).$$

- Quality of approximation?

$$\mathrm{W}_2(\nabla K(\psi), \widetilde{\nabla K}(\psi)) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?$$

Motivations

2. *Linearized Optimal Transport*

- Recall that

$$\boxed{\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id} \text{ where } \phi_\mu \in \arg\min_{\phi \text{ convex}} \langle \phi | \rho \rangle + \langle \phi^* | \mu \rangle.$$

- For a fixed $\rho \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$, distance in $\mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$ is

$$\boxed{\mathrm{W}_{2,\rho}(\mu, \nu) := \|\nabla\phi_\mu - \nabla\phi_\nu\|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}.}$$

$\rightarrow$ *Linearized Optimal Transport* distance (Wang et al., 2013)

- *Linearized/generalized* geodesics:

$$\boxed{\mu_t = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \rho}$$

- Approximation with $\tilde{\mu_t} = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \tilde{\rho}$.

$$\boxed{\mathrm{W}_2(\mu_t, \tilde{\mu_t}) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?}$$

Motivations

2. *Linearized Optimal Transport*

- Recall that

$$\boxed{\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id} \text{ where } \phi_\mu \in \arg\min_{\phi \text{ convex}} \langle \phi | \rho \rangle + \langle \phi^* | \mu \rangle.$$

- For a fixed $\rho \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$, distance in $\mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$ is

$$\mathrm{W}_{2,\rho}(\mu, \nu) := \|\nabla\phi_\mu - \nabla\phi_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)}.$$

  $\rightarrow$ *Linearized Optimal Transport* distance (Wang et al., 2013)

- *Linearized/generalized* geodesics:

$$\mu_t = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \rho$$

- Approximation with $\tilde{\mu_t} = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \tilde{\rho}$.

$$\mathrm{W}_2(\mu_t, \tilde{\mu_t}) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?$$

Motivations

2. *Linearized Optimal Transport*

- Recall that

$$\boxed{\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id} \text{ where } \phi_\mu \in \arg\min_{\phi \text{ convex}} \langle\phi|\rho\rangle + \langle\phi^*|\mu\rangle.$$

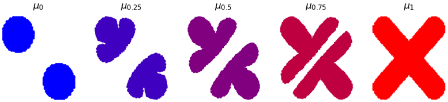
- For a fixed $\rho \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$, distance in $\mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$ is

$$\mathrm{W}_{2,\rho}(\mu,\nu) := \|\nabla\phi_\mu - \nabla\phi_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)}.$$

$\rightarrow$ *Linearized Optimal Transport* distance (Wang et al., 2013)

- *Linearized/generalized* geodesics:

$$\mu_t = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \rho$$



- Approximation with $\tilde{\mu_t} = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \tilde{\rho}$.

$$\mathrm{W}_2(\mu_t, \tilde{\mu_t}) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?$$

Motivations

2. *Linearized Optimal Transport*

- Recall that

$$\boxed{\exp_\rho^{-1}(\mu) = \nabla\phi_\mu - id} \text{ where } \phi_\mu \in \arg\min_{\phi \text{ convex}} \langle \phi | \rho \rangle + \langle \phi^* | \mu \rangle.$$

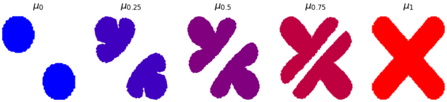
- For a fixed $\rho \in \mathcal{P}_2^{a.c.}(\mathbb{R}^d)$, distance in $\mathcal{T}_\rho\mathcal{P}_2(\mathbb{R}^d)$ is

$$\mathrm{W}_{2,\rho}(\mu,\nu) := \|\nabla\phi_\mu - \nabla\phi_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)}.$$

$\rightarrow$ *Linearized Optimal Transport* distance (Wang et al., 2013)

- *Linearized/generalized* geodesics:

$$\mu_t = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \rho$$



- Approximation with $\tilde{\mu}_t = ((1-t)\nabla\phi_{\mu_0} + t\nabla\phi_{\mu_1})_\# \tilde{\rho}$.

$$\mathrm{W}_2(\mu_t, \tilde{\mu}_t) \text{ vs. } \mathrm{W}_2(\rho, \tilde{\rho})?$$

Motivations

3. Generative modelling with *Input Convex Neural Networks*

Input Convex Neural Network (Amos et al., 2017):
Neural network $\phi_\theta : \mathbb{R}^d \to \mathbb{R}$, parameterized by $\theta \in \Theta$, whose architecture constraints $x \mapsto \phi_\theta(x)$ to be convex.

- Generative modelling with ICNNs:

$$\min_\theta \mathcal{L}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \rho, \mu).$$

- In practice $\not\rho, \not\mu \to$ statistical approximations $\hat{\rho}, \hat{\mu}$:

$$\min_\theta \hat{\mathcal{L}}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \hat{\rho}, \hat{\mu}).$$

- Empirical risk minimization bounds?

$$\left|\hat{\mathcal{L}}(\theta) - \mathcal{L}(\theta)\right| \text{ vs. } \mathrm{W}_2(\rho, \hat{\rho}) + \mathrm{W}_2(\mu, \hat{\mu})?$$

Motivations

3. Generative modelling with *Input Convex Neural Networks*

Input Convex Neural Network (Amos et al., 2017):
Neural network $\phi_\theta : \mathbb{R}^d \to \mathbb{R}$, parameterized by $\theta \in \Theta$, whose architecture constraints $x \mapsto \phi_\theta(x)$ to be convex.

- Generative modelling with ICNNs:

$$\min_\theta \mathcal{L}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \rho, \mu).$$

- In practice $\not\rho, \not\mu \to$ statistical approximations $\hat{\rho}, \hat{\mu}$:

$$\min_\theta \hat{\mathcal{L}}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \hat{\rho}, \hat{\mu}).$$

- Empirical risk minimization bounds?

$$\left|\hat{\mathcal{L}}(\theta) - \mathcal{L}(\theta)\right| \text{ vs. } \mathrm{W}_2(\rho, \hat{\rho}) + \mathrm{W}_2(\mu, \hat{\mu})?$$

Motivations

3. Generative modelling with *Input Convex Neural Networks*

Input Convex Neural Network (Amos et al., 2017):
Neural network $\phi_\theta : \mathbb{R}^d \to \mathbb{R}$, parameterized by $\theta \in \Theta$, whose architecture constraints $x \mapsto \phi_\theta(x)$ to be convex.

- Generative modelling with ICNNs:

$$\min_\theta \mathcal{L}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \rho, \mu).$$

- In practice $\cancel{\rho}, \cancel{\mu} \to$ statistical approximations $\hat{\rho}, \hat{\mu}$:

$$\min_\theta \hat{\mathcal{L}}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \hat{\rho}, \hat{\mu}).$$

- Empirical risk minimization bounds?

$$\left|\hat{\mathcal{L}}(\theta) - \mathcal{L}(\theta)\right| \text{ vs. } \mathrm{W}_2(\rho, \hat{\rho}) + \mathrm{W}_2(\mu, \hat{\mu})?$$

Motivations

3. Generative modelling with *Input Convex Neural Networks*

Input Convex Neural Network (Amos et al., 2017):
Neural network $\phi_\theta : \mathbb{R}^d \to \mathbb{R}$, parameterized by $\theta \in \Theta$, whose architecture constraints $x \mapsto \phi_\theta(x)$ to be convex.

- Generative modelling with ICNNs:

$$\min_\theta \mathcal{L}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \rho, \mu).$$

- In practice ~~ρ~~, ~~μ~~ $\to$ statistical approximations $\hat{\rho}, \hat{\mu}$:

$$\min_\theta \hat{\mathcal{L}}(\theta) \approx \mathrm{W}_2((\nabla\phi_\theta)_\# \hat{\rho}, \hat{\mu}).$$

- Empirical risk minimization bounds?

$$\left|\hat{\mathcal{L}}(\theta) - \mathcal{L}(\theta)\right| \text{ vs. } \mathrm{W}_2(\rho, \hat{\rho}) + \mathrm{W}_2(\mu, \hat{\mu})?$$

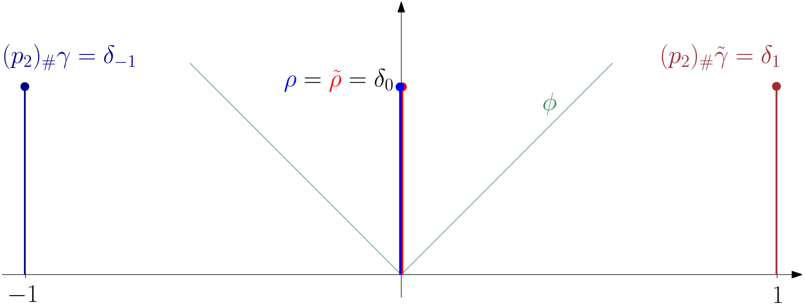
A positive result

Proposition:
Let $\alpha \in (0,1)$ and let $\phi \in \mathcal{C}^{1,\alpha}(\mathbb{R}^d)$ convex. Then for any $\rho, \tilde{\rho} \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\mathrm{W}_2((\nabla\phi)_{\#}\rho, (\nabla\phi)_{\#}\tilde{\rho}) \leq \|\nabla\phi\|_{\mathcal{C}^{0,\alpha}} \mathrm{W}_2(\rho, \tilde{\rho})^{\alpha}.$$

Negative results

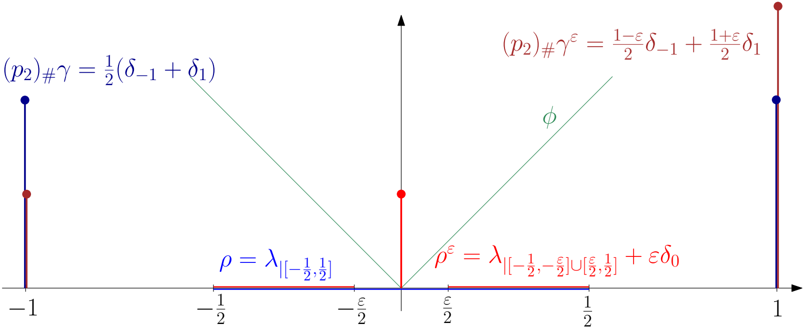
- No assumption on $\rho, \tilde{\rho}$:



$$\mathrm{W}_2((p_2)_\# \gamma, (p_2)_\# \tilde{\gamma}) = 2 \text{ while } \mathrm{W}_2(\rho, \tilde{\rho}) = 0.$$

Negative results

- Assume ρ is absolutely continuous and $\rho \leq M < +\infty$:



$$\mathrm{W}_2((p_2)_\# \gamma, (p_2)_\# \gamma^\varepsilon) \sim \mathrm{W}_2(\rho, \rho^\varepsilon)^{1/3}.$$

Main result

Assumptions:

- Let $R > 0$ and let $\Omega = B(0, R) \subset \mathbb{R}^d$.
- Let $\phi : \Omega \to \mathbb{R}$ convex and R-Lipschitz continuous.
- Let $M \in (0, +\infty)$.

Theorem:

- For any $\rho \in \mathcal{P}_{a.c.}(\Omega)$ s.t. $\rho \leq M$,
- For any $\tilde{\rho} \in \mathcal{P}(\Omega)$ and $\tilde{\gamma} \in \mathcal{P}(\Omega \times \Omega)$ s.t. $(p_1)_\# \tilde{\gamma} = \tilde{\rho}$ and $\mathrm{spt}(\tilde{\gamma}) \subset \partial\phi$,

$$\boxed{\mathrm{W}_2((\nabla\phi)_\#\rho, (p_2)_\#\tilde{\gamma}) \leq C(d, M, R)\mathrm{W}_2(\rho, \tilde{\rho})^{1/3},}$$

where $C(d, M, R) \sim d^2 2^{8(d+1)}(1 + \beta_d)(1 + M)(1 + R)^{4+d}$.

Main result

Assumptions:

- Let $R > 0$ and let $\Omega = B(0, R) \subset \mathbb{R}^d$.
- Let $\phi : \Omega \to \mathbb{R}$ convex and R-Lipschitz continuous.
- Let $M \in (0, +\infty)$.

Theorem:

- For any $\rho \in \mathcal{P}_{a.c.}(\Omega)$ s.t. $\rho \leq M$,
- For any $\tilde{\rho} \in \mathcal{P}(\Omega)$ and $\tilde{\gamma} \in \mathcal{P}(\Omega \times \Omega)$ s.t. $(p_1)_{\#}\tilde{\gamma} = \tilde{\rho}$ and $\mathrm{spt}(\tilde{\gamma}) \subset \partial\phi$,

$$\boxed{\mathrm{W}_2((\nabla\phi)_{\#}\rho, (p_2)_{\#}\tilde{\gamma}) \leq C(d, M, R)\mathrm{W}_2(\rho, \tilde{\rho})^{\mathbf{1/3}},}$$

where $C(d, M, R) \sim d^2 2^{8(d+1)}(1 + \beta_d)(1 + M)(1 + R)^{4+d}$.

Key ingredient

Covering number of near-singular sets of convex functions

$$\Sigma_{\eta,\alpha} := \{x \in \Omega \mid \mathrm{diam}(\partial\phi(B(x,\eta)) \geq \alpha\}.$$

Theorem: For all $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{R^{d-1}}{\eta^{d-1}} \frac{\mathrm{Lip}(\phi)}{\alpha}.$$

Proof in dimension 1: $\Sigma_{\eta,\alpha} := \{x \in [-R, R] \mid \phi'(x+\eta) - \phi'(x-\eta) \geq \alpha\}$.

Let $\{x_i\}_{1 \leq i \leq N}$ be an ordered η-packing of $\Sigma_{\eta,\alpha}$. Then

$$\begin{aligned} N\alpha &\leq \sum_{i=1}^{N} \phi'(x_i+\eta) - \phi'(x_i-\eta) \\ &= \phi'(x_N+\eta) - \phi'(x_1-\eta) + \sum_{i=1}^{N-1} \underbrace{\phi'(x_i+\eta) - \phi'(x_{i+1}-\eta)}_{\leq 0 \text{ since } x_i+\eta < x_{i+1}-\eta \text{ and } \phi' \nearrow} \\ &\leq \phi'(x_N+\eta) - \phi'(x_1-\eta) \\ &\leq 2\mathrm{Lip}(\phi). \end{aligned}$$

Key ingredient

Covering number of near-singular sets of convex functions

$$\Sigma_{\eta,\alpha} := \{x \in \Omega \mid \mathrm{diam}(\partial\phi(B(x,\eta)) \geq \alpha\}.$$

Theorem: For all $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{R^{d-1}}{\eta^{d-1}} \frac{\mathrm{Lip}(\phi)}{\alpha}.$$

Proof in dimension 1: $\Sigma_{\eta,\alpha} := \{x \in [-R, R] \mid \phi'(x+\eta) - \phi'(x-\eta) \geq \alpha\}$.

Let $\{x_i\}_{1\leq i\leq N}$ be an ordered η-packing of $\Sigma_{\eta,\alpha}$. Then

$$\begin{aligned} N\alpha &\leq \sum_{i=1}^{N} \phi'(x_i+\eta) - \phi'(x_i-\eta) \\ &= \phi'(x_N+\eta) - \phi'(x_1-\eta) + \sum_{i=1}^{N-1} \underbrace{\phi'(x_i+\eta) - \phi'(x_{i+1}-\eta)}_{\leq 0 \text{ since } x_i+\eta < x_{i+1}-\eta \text{ and } \phi' \nearrow} \\ &\leq \phi'(x_N+\eta) - \phi'(x_1-\eta) \\ &\leq 2\mathrm{Lip}(\phi). \end{aligned}$$

Key ingredient

Covering number of near-singular sets of convex functions

$$\Sigma_{\eta,\alpha} := \{x \in \Omega \mid \mathrm{diam}(\partial\phi(B(x,\eta)) \geq \alpha\}.$$

Theorem: For all $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{R^{d-1}}{\eta^{d-1}} \frac{\mathrm{Lip}(\phi)}{\alpha}.$$

Proof in dimension 1**:** $\Sigma_{\eta,\alpha} := \{x \in [-R, R] \mid \phi'(x+\eta) - \phi'(x-\eta) \geq \alpha\}$.

Let $\{x_i\}_{1\leq i\leq N}$ be an ordered η-packing of $\Sigma_{\eta,\alpha}$. Then

$$\begin{aligned} N\alpha &\leq \sum_{i=1}^{N} \phi'(x_i+\eta) - \phi'(x_i-\eta) \\ &= \phi'(x_N+\eta) - \phi'(x_1-\eta) + \sum_{i=1}^{N-1} \underbrace{\phi'(x_i+\eta) - \phi'(x_{i+1}-\eta)}_{\leq 0 \text{ since } x_i+\eta < x_{i+1}-\eta \text{ and } \phi' \nearrow} \\ &\leq \phi'(x_N+\eta) - \phi'(x_1-\eta) \\ &\leq 2\mathrm{Lip}(\phi). \end{aligned}$$

Key ingredient

Covering number of near-singular sets of convex functions

$$\Sigma_{\eta,\alpha} := \{x \in \Omega \mid \mathrm{diam}(\partial\phi(B(x,\eta)) \geq \alpha\}.$$

Theorem: For all $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{R^{d-1}}{\eta^{d-1}} \frac{\mathrm{Lip}(\phi)}{\alpha}.$$

Proof in dimension 1: $\Sigma_{\eta,\alpha} := \{x \in [-R, R] \mid \phi'(x+\eta) - \phi'(x-\eta) \geq \alpha\}$.

Let $\{x_i\}_{1\leq i\leq N}$ be an ordered η-packing of $\Sigma_{\eta,\alpha}$. Then

$$N\alpha \leq \sum_{i=1}^{N} \phi'(x_i+\eta) - \phi'(x_i-\eta)$$

$$= \phi'(x_N+\eta) - \phi'(x_1-\eta) + \sum_{i=1}^{N-1} \underbrace{\phi'(x_i+\eta) - \phi'(x_{i+1}-\eta)}_{\leq 0 \text{ since } x_i+\eta < x_{i+1}-\eta \text{ and } \phi' \nearrow}$$

$$\leq \phi'(x_N+\eta) - \phi'(x_1-\eta)$$

$$\leq 2\mathrm{Lip}(\phi).$$

Key ingredient

Covering number of near-singular sets of convex functions

$$\Sigma_{\eta,\alpha} := \{x \in \Omega \mid \mathrm{diam}(\partial\phi(B(x,\eta)) \geq \alpha\}.$$

Theorem: For all $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{R^{d-1}}{\eta^{d-1}} \frac{\mathrm{Lip}(\phi)}{\alpha}.$$

Proof in dimension 1**:** $\Sigma_{\eta,\alpha} := \{x \in [-R, R] \mid \phi'(x+\eta) - \phi'(x-\eta) \geq \alpha\}$.

Let $\{x_i\}_{1\leq i\leq N}$ be an ordered η-packing of $\Sigma_{\eta,\alpha}$. Then

$$\begin{aligned} N\alpha &\leq \sum_{i=1}^{N} \phi'(x_i+\eta) - \phi'(x_i-\eta) \\ &= \phi'(x_N+\eta) - \phi'(x_1-\eta) + \sum_{i=1}^{N-1} \underbrace{\phi'(x_i+\eta) - \phi'(x_{i+1}-\eta)}_{\leq 0 \text{ since } x_i+\eta<x_{i+1}-\eta \text{ and } \phi' \nearrow} \\ &\leq \phi'(x_N+\eta) - \phi'(x_1-\eta) \\ &\leq 2\mathrm{Lip}(\phi). \end{aligned}$$

Key ingredient

Covering number of near-singular sets of convex functions

$$\Sigma_{\eta,\alpha} := \{x \in \Omega \mid \mathrm{diam}(\partial\phi(B(x,\eta)) \geq \alpha\}.$$

Theorem: For all $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{R^{d-1}}{\eta^{d-1}} \frac{\mathrm{Lip}(\phi)}{\alpha}.$$

Proof in dimension 1**:** $\Sigma_{\eta,\alpha} := \{x \in [-R,R] \mid \phi'(x+\eta) - \phi'(x-\eta) \geq \alpha\}$.

Let $\{x_i\}_{1\leq i\leq N}$ be an ordered η-packing of $\Sigma_{\eta,\alpha}$. Then

$$\begin{aligned}
N\alpha &\leq \sum_{i=1}^{N} \phi'(x_i+\eta) - \phi'(x_i-\eta) \\
&= \phi'(x_N+\eta) - \phi'(x_1-\eta) + \sum_{i=1}^{N-1} \underbrace{\phi'(x_i+\eta) - \phi'(x_{i+1}-\eta)}_{\leq 0 \text{ since } x_i+\eta < x_{i+1}-\eta \text{ and } \phi' \nearrow} \\
&\leq \phi'(x_N+\eta) - \phi'(x_1-\eta) \\
&\leq 2\mathrm{Lip}(\phi).
\end{aligned}$$

Key ingredient

Covering number of near-singular sets of convex functions

$$\Sigma_{\eta,\alpha} := \{x \in \Omega \mid \mathrm{diam}(\partial\phi(B(x,\eta)) \geq \alpha\}.$$

Theorem: For all $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{R^{d-1}}{\eta^{d-1}} \frac{\mathrm{Lip}(\phi)}{\alpha}.$$

Proof in dimension 1: $\Sigma_{\eta,\alpha} := \{x \in [-R, R] \mid \phi'(x+\eta) - \phi'(x-\eta) \geq \alpha\}$.

Let $\{x_i\}_{1\leq i\leq N}$ be an ordered η-packing of $\Sigma_{\eta,\alpha}$. Then

$$\begin{aligned}
N\alpha &\leq \sum_{i=1}^{N} \phi'(x_i+\eta) - \phi'(x_i-\eta) \\
&= \phi'(x_N+\eta) - \phi'(x_1-\eta) + \sum_{i=1}^{N-1} \underbrace{\phi'(x_i+\eta) - \phi'(x_{i+1}-\eta)}_{\leq 0 \text{ since } x_i+\eta<x_{i+1}-\eta \text{ and } \phi' \nearrow} \\
&\leq \phi'(x_N+\eta) - \phi'(x_1-\eta) \\
&\leq 2\mathrm{Lip}(\phi).
\end{aligned}$$

Comparison

On the singularities of convex functions

$$\Sigma_{0,\alpha} = \{x \in \Omega \mid \mathrm{diam}(\partial\phi(x)) \geq \alpha\}.$$

We recover that $\dim_{\mathcal{H}}(\Sigma_{0,\alpha}) \leq d - 1$ and

$$\mathcal{H}^{d-1}(\Sigma_{0,\alpha}) \leq C(d) \frac{R^{d-1}\mathrm{Lip}(\phi)}{\alpha}.$$

Theorem (Alberti, Ambrosio, Cannarsa, 1992):
Let $k \in \{1, \ldots, d\}$. The set

$$\Sigma^k := \{x \in \Omega \mid \dim_{\mathcal{H}}(\partial\phi(x)) \geq k\}$$

is countably $\mathcal{H}^{d-k}$-rectifiable. It satisfies

$$\int_{\Sigma^k} \mathcal{H}^k(\partial\phi(x)) \mathrm{d}\mathcal{H}^{d-k}(x) \leq C(d)(\mathrm{Lip}(\phi) + 2R)^d.$$

This yields

$$\mathcal{H}^{d-1}(\Sigma_{0,\alpha}) \leq C(d) \frac{(\mathrm{Lip}(\phi) + 2R)^d}{\alpha}.$$

Comparison

On the singularities of convex functions

$$\Sigma_{0,\alpha} = \{x \in \Omega \mid \mathrm{diam}(\partial\phi(x)) \geq \alpha\}.$$

We recover that $\dim_{\mathcal{H}}(\Sigma_{0,\alpha}) \leq d - 1$ and

$$\mathcal{H}^{d-1}(\Sigma_{0,\alpha}) \leq C(d)\frac{R^{d-1}\mathrm{Lip}(\phi)}{\alpha}.$$

Theorem (Alberti, Ambrosio, Cannarsa, 1992):
Let $k \in \{1, \dots, d\}$. The set

$$\Sigma^k := \{x \in \Omega \mid \dim_{\mathcal{H}}(\partial\phi(x)) \geq k\}$$

is countably $\mathcal{H}^{d-k}$-rectifiable. It satisfies

$$\int_{\Sigma^k} \mathcal{H}^k(\partial\phi(x))\mathrm{d}\mathcal{H}^{d-k}(x) \leq C(d)(\mathrm{Lip}(\phi) + 2R)^d.$$

This yields

$$\mathcal{H}^{d-1}(\Sigma_{0,\alpha}) \leq C(d)\frac{(\mathrm{Lip}(\phi) + 2R)^d}{\alpha}.$$

Comparison

On the singularities of convex functions

$$\Sigma_{0,\alpha} = \{x \in \Omega \mid \mathrm{diam}(\partial\phi(x)) \geq \alpha\}.$$

We recover that $\dim_{\mathcal{H}}(\Sigma_{0,\alpha}) \leq d-1$ and

$$\mathcal{H}^{d-1}(\Sigma_{0,\alpha}) \leq C(d)\frac{R^{d-1}\mathrm{Lip}(\phi)}{\alpha}.$$

Theorem (Alberti, Ambrosio, Cannarsa, 1992):
Let $k \in \{1,\dots,d\}$. The set

$$\Sigma^k := \{x \in \Omega \mid \dim_{\mathcal{H}}(\partial\phi(x)) \geq k\}$$

is countably $\mathcal{H}^{d-k}$-rectifiable. It satisfies

$$\int_{\Sigma^k} \mathcal{H}^k(\partial\phi(x))\mathrm{d}\mathcal{H}^{d-k}(x) \leq C(d)(\mathrm{Lip}(\phi)+2R)^d.$$

This yields

$$\mathcal{H}^{d-1}(\Sigma_{0,\alpha}) \leq C(d)\frac{(\mathrm{Lip}(\phi)+2R)^d}{\alpha}.$$

Main result

General setting

Assumptions:

- Let $R > 0$ and let $\Omega = B(0, R) \subset \mathbb{R}^d$.
- Let $p \geq 2$ and $c(x, y) = \|x - y\|^p$. Let $\varphi \in \mathcal{C}(\Omega)$ satisfying $\varphi = (\varphi^c)^{\bar{c}}$. Denote
$$T_\varphi : x \mapsto x - (\nabla \|\cdot\|^p)^{-1}(\nabla\varphi(x)).$$
- Let $M \in (0, +\infty)$.

Theorem:

- For any $\rho \in \mathcal{P}_{a.c.}(\Omega)$ s.t. $\rho \leq M$,
- For any $\tilde{\rho} \in \mathcal{P}(\Omega)$ and $\tilde{\gamma} \in \mathcal{P}(\Omega \times \Omega)$ s.t. $(p_1)_\#\tilde{\gamma} = \tilde{\rho}$ and $\mathrm{spt}(\tilde{\gamma}) \subset \partial^c\varphi$,
- For any $q \in (p - 1, \infty)$ and $r \in (1, \infty]$

$$\mathrm{W}_q((T_\varphi)_\#\rho, (p_2)_\#\tilde{\gamma}) \leq C(d, q, p, M, R)\mathrm{W}_r(\rho, \tilde{\rho})^{\frac{r}{q(r+1)}},$$

where $C(d, q, p, M, R) \sim 2^{8(d+1)}p^3 \left(\frac{q}{q-p+1}\right)^{1/q} d^2(1 + \beta_d)(1 + M_\rho)(1 + R)^{2+p+d}$.

Thank you for your attention!

Main result

General setting

Assumptions:

- Let $R > 0$ and let $\Omega = B(0, R) \subset \mathbb{R}^d$.
- Let $p \geq 2$ and $c(x, y) = \|x - y\|^p$. Let $\varphi \in \mathcal{C}(\Omega)$ satisfying $\varphi = (\varphi^c)^{\bar{c}}$. Denote

$$T_\varphi : x \mapsto x - (\nabla \|\cdot\|^p)^{-1}(\nabla \varphi(x)).$$

- Let $M \in (0, +\infty)$.

Theorem:

- For any $\rho \in \mathcal{P}_{a.c.}(\Omega)$ s.t. $\rho \leq M$,
- For any $\tilde{\rho} \in \mathcal{P}(\Omega)$ and $\tilde{\gamma} \in \mathcal{P}(\Omega \times \Omega)$ s.t. $(p_1)_\# \tilde{\gamma} = \tilde{\rho}$ and $\mathrm{spt}(\tilde{\gamma}) \subset \partial^c \varphi$,
- For any $q \in (p - 1, \infty)$ and $r \in (1, \infty]$,

$$\boxed{\mathrm{W}_q((T_\varphi)_\# \rho, (p_2)_\# \tilde{\gamma}) \leq C(d, q, p, M, R) \mathrm{W}_r(\rho, \tilde{\rho})^{\frac{r}{q(r+1)}},}$$

where $C(d, q, p, M, R) \sim 2^{8(d+1)} p^3 \left(\frac{q}{q-p+1}\right)^{1/q} d^2 (1 + \beta_d)(1 + M_\rho)(1 + R)^{2+p+d}$.

Thank you for your attention!

Main result

General setting

Assumptions:

- Let $R > 0$ and let $\Omega = B(0, R) \subset \mathbb{R}^d$.
- Let $p \geq 2$ and $c(x, y) = \|x - y\|^p$. Let $\varphi \in \mathcal{C}(\Omega)$ satisfying $\varphi = (\varphi^c)^{\bar{c}}$. Denote
$$T_\varphi : x \mapsto x - (\nabla \|\cdot\|^p)^{-1}(\nabla \varphi(x)).$$
- Let $M \in (0, +\infty)$.

Theorem:

- For any $\rho \in \mathcal{P}_{a.c.}(\Omega)$ s.t. $\rho \leq M$,
- For any $\tilde{\rho} \in \mathcal{P}(\Omega)$ and $\tilde{\gamma} \in \mathcal{P}(\Omega \times \Omega)$ s.t. $(p_1)_\# \tilde{\gamma} = \tilde{\rho}$ and $\mathrm{spt}(\tilde{\gamma}) \subset \partial^c \varphi$,
- For any $q \in (p-1, \infty)$ and $r \in (1, \infty]$,

$$\boxed{\mathrm{W}_q((T_\varphi)_\# \rho, (p_2)_\# \tilde{\gamma}) \leq C(d, q, p, M, R) \mathrm{W}_r(\rho, \tilde{\rho})^{\frac{r}{q(r+1)}}},$$

where $C(d, q, p, M, R) \sim 2^{8(d+1)} p^3 \left(\frac{q}{q-p+1}\right)^{1/q} d^2 (1 + \beta_d)(1 + M_\rho)(1 + R)^{2+p+d}$.

Thank you for your attention!