Quantitative Stability of the Pushforward Operation by an Optimal Transport Map

Alex Delalande

Joint work with Guillaume Carlier and Quentin Mérigot

November 2023

Problem statement

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

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

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

Under what conditions on $\phi, \rho, \tilde{\rho}$ and for which C, α do we have

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

Remark: whenever ϕ is differentiable $\rho-$ and $\tilde{\rho}-$a.e.,

$$\gamma = (\mathrm{id}, \nabla\phi)_\# \rho, \quad \tilde{\gamma} = (\mathrm{id}, \nabla\phi)_\# \tilde{\rho},$$

$$\mathrm{W}_2((p_2)_\# \gamma, (p_2)_\# \tilde{\gamma}) = \mathrm{W}_2((\nabla\phi)_\# \rho, (\nabla\phi)_\# \tilde{\rho}).$$

Problem statement

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

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

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

Under what conditions on $\phi, \rho, \tilde{\rho}$ and for which C, α do we have

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

Remark: whenever ϕ is differentiable $\rho-$ and $\tilde{\rho}-$a.e.,

$$\gamma = (\mathrm{id}, \nabla \phi)_\# \rho, \quad \tilde{\gamma} = (\mathrm{id}, \nabla \phi)_\# \tilde{\rho},$$

$$\mathrm{W}_2((p_2)_\# \gamma, (p_2)_\# \tilde{\gamma}) = \mathrm{W}_2((\nabla \phi)_\# \rho, (\nabla \phi)_\# \tilde{\rho}).$$

Motivations

1. **Resolution of Kantorovich dual:**

$$\min_{\psi} \underbrace{\langle \psi^* | \rho \rangle}_{:=K(\psi)} + \langle \psi | \mu \rangle.$$

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

2. **Barycenters in *Linearized OT*:**

$$\mathrm{Bar}_\rho((\mu_i)_{1 \le i \le N}) = \left(\frac{1}{N} \sum_i \nabla \phi_{\rho \to \mu_i} \right)_{\#} \rho.$$

3. **Generative modelling with an *ICNN* ϕ_θ:**

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

Motivations

1. **Resolution of Kantorovich dual:**

$$\min_{\psi} \underbrace{\langle \psi^* | \rho \rangle}_{:= K(\psi)} + \langle \psi | \mu \rangle.$$

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

2. **Barycenters in *Linearized OT*:**

$$\boxed{\mathrm{Bar}_\rho((\mu_i)_{1 \le i \le N}) = \left(\frac{1}{N} \sum_i \nabla \phi_{\rho \to \mu_i} \right)_{\#} \rho.}$$

3. **Generative modelling with an *ICNN* ϕ_θ:**

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

Motivations

1. **Resolution of Kantorovich dual:**

$$\min_{\psi} \underbrace{\langle \psi^* | \rho \rangle}_{:= K(\psi)} + \langle \psi | \mu \rangle.$$

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

2. **Barycenters in *Linearized OT*:**

$$\mathrm{Bar}_{\rho}((\mu_i)_{1 \le i \le N}) = \left(\frac{1}{N} \sum_i \nabla \phi_{\rho \to \mu_i} \right)_{\#} \rho.$$

3. **Generative modelling with an *ICNN* ϕ_θ:**

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

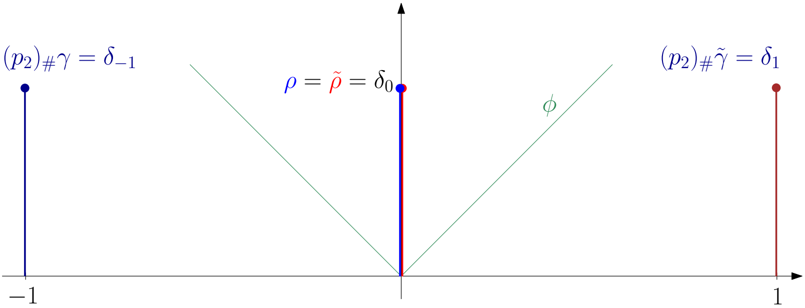
A positive result

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

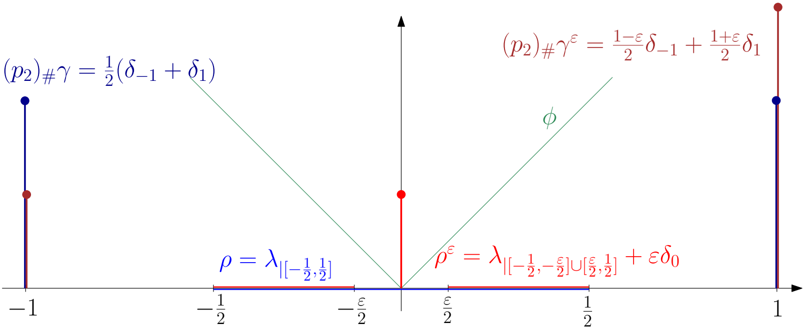
- No assumption on $\rho, \tilde{\rho}$:



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

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 $\mathrm{spt}(\tilde{\gamma}) \subset \partial\phi$,

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

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 $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{d^2 R^{d-1}}{\alpha \eta^{d-1}}.$$

In particular,

$$\boxed{\int_\Omega \mathrm{diam}(\partial\phi(B(x,\eta)))^2 \mathrm{d}x \leq C(d,R)\eta,}$$

where $C(d,R) \sim 2^{3d} d^2 \beta_d R^{d+1}$.

Remark: also entails that $\dim_{\mathcal{H}}(\Sigma_{0,\alpha}) \leq d-1$ and

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

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 $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{d^2 R^{d-1}}{\alpha \eta^{d-1}}.$$

In particular,

$$\boxed{\int_\Omega \mathrm{diam}(\partial\phi(B(x,\eta)))^2 \mathrm{d}x \leq C(d,R)\eta,}$$

where $C(d,R) \sim 2^{3d} d^2 \beta_d R^{d+1}$.

Remark: also entails that $\dim_{\mathcal{H}}(\Sigma_{0,\alpha}) \leq d-1$ and

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

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 $\alpha, \eta > 0$,

$$\mathcal{N}(\Sigma_{\eta,\alpha}, \eta) \lesssim \frac{d^2 R^{d-1}}{\alpha \eta^{d-1}}.$$

In particular,

$$\boxed{\int_\Omega \mathrm{diam}(\partial\phi(B(x,\eta)))^2 \mathrm{d}x \leq C(d,R)\eta,}$$

where $C(d,R) \sim 2^{3d} d^2 \beta_d R^{d+1}$.

Remark: also entails that $\dim_{\mathcal{H}}(\Sigma_{0,\alpha}) \leq d-1$ and

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

Comparison

On the singularities of convex functions

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

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!