Quantitative Stability in Quadratic Optimal Transport

Alex Delalande

EPFL

CIRM Workshop *PDE & Probability in interaction: functional inequalities, optimal transport and particle systems*

January 26, 2024

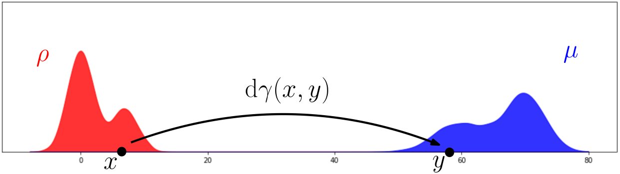
Introduction

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

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

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

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



Introduction

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

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

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

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

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

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

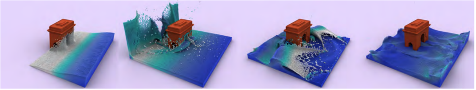
- 2−***Wasserstein space***: $(\mathcal{P}_2(\mathbb{R}^d), \mathrm{W}_2)$.

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

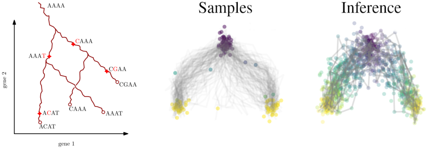
Introduction

- **Strong physical modeling power:**

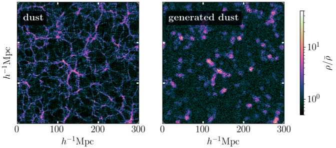
Euler equations: (de Goes et al., 2015)



Trajectory inference for single cell RNA-sequencing data: (Forrow et al., 2021; Chizat et al., 2022)



Cosmology: (Nikakhtar et al., 2023)



Quantum chemistry, **meteorology**, **economics**, **image processing**, **machine learning**...

Introduction

Is the quadratic optimal transport problem $\inf_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle$ ***well-posed*?**

1. **Existence of optimal γ? Verified.**
   (Kantorovich, 1942; Kellerer, 1984)

2. **Uniqueness of optimal γ? Not verified in general.**
   But well-studied and verified in many *particular* cases.

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

For μ a.c. and under additional assumptions, optimal ϕ solution of the **Monge-Ampère equation**:

$$\forall x \in \Omega, \quad \det(\mathrm{D}^2\phi(x))\mu(\nabla\phi(x)) = \rho(x).$$

Introduction

Is the quadratic optimal transport problem $\inf_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle$ ***well-posed*?**

1. **Existence of optimal γ? Verified.**
   (Kantorovich, 1942; Kellerer, 1984)

2. **Uniqueness of optimal γ? Not verified in general.**
   But well-studied and verified in many *particular* cases.

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

For μ a.c. and under additional assumptions, optimal ϕ solution of the **Monge-Ampère equation**:

$$\forall x \in \Omega, \quad \det(\mathrm{D}^2 \phi(x)) \mu(\nabla \phi(x)) = \rho(x).$$

Introduction

Is the quadratic optimal transport problem $\inf_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle$ ***well-posed*****?**

3. Stability of optimal γ w.r.t. problem data? Verified.

Theorem (Villani, 2008): If $\rho_n \rightharpoonup \rho$, $\mu_n \rightharpoonup \mu$, then for $\gamma_n \in \Gamma(\rho_n, \mu_n)$ **optimal**, up to a subsequence, $\gamma_n \rightharpoonup \gamma \in \Gamma(\rho, \mu)$ **optimal**.

Example:

What is the *rate* of the convergence $\gamma_n \rightharpoonup \gamma$ in terms of the *rates* of convergence of $\rho_n \rightharpoonup \rho$ and $\mu_n \rightharpoonup \mu$?

$\rightarrow$ **Quantitative stability of optimal transport solutions?**

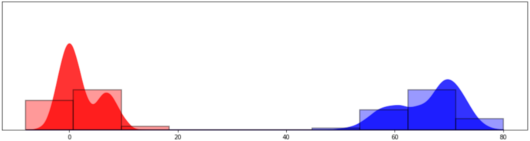
Introduction

Is the quadratic optimal transport problem $\inf_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle$ ***well-posed*?**

3. Stability of optimal γ w.r.t. problem data? Verified.

Theorem (Villani, 2008): If $\rho_n \rightharpoonup \rho$, $\mu_n \rightharpoonup \mu$, then for $\gamma_n \in \Gamma(\rho_n, \mu_n)$ **optimal**, up to a subsequence, $\gamma_n \rightharpoonup \gamma \in \Gamma(\rho, \mu)$ **optimal**.

Example:



What is the *rate* of the convergence $\gamma_n \rightharpoonup \gamma$ in terms of the *rates* of convergence of $\rho_n \rightharpoonup \rho$ and $\mu_n \rightharpoonup \mu$?

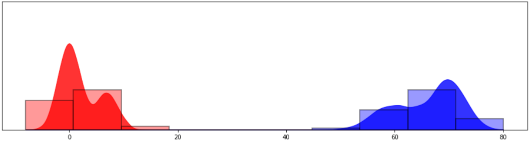
Introduction

Is the quadratic optimal transport problem $\inf_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle$ ***well-posed*?**

3. Stability of optimal γ w.r.t. problem data? Verified.

Theorem (Villani, 2008): If $\rho_n \rightharpoonup \rho$, $\mu_n \rightharpoonup \mu$, then for $\gamma_n \in \Gamma(\rho_n, \mu_n)$ **optimal**, up to a subsequence, $\gamma_n \rightharpoonup \gamma \in \Gamma(\rho, \mu)$ **optimal**.

Example:



What is the *rate* of the convergence $\gamma_n \rightharpoonup \gamma$ in terms of the *rates* of convergence of $\rho_n \rightharpoonup \rho$ and $\mu_n \rightharpoonup \mu$?

Introduction

- **Quantitative stability of optimal transport solutions w.r.t. problem data?**

1. Beyond the case $\Omega = \mathbb{R}$, no general quantitative stability result.
2. Convex optimization viewpoint: $\min_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle$.
   $\rightarrow$ **Strong/Uniform convexity?** *Linear program.*
3. Elliptic PDE viewpoint: $\det(\mathrm{D}^2\phi)\mu(\nabla\phi) = \rho.$
   $\rightarrow$ **Uniform ellipticity?** *Degenerate elliptic equation.*

Using the unconstrained dual problem, quantitative stability guarantees can be obtained.

Introduction

- **Quantitative stability of optimal transport solutions w.r.t. problem data?**

1. Beyond the case $\Omega = \mathbb{R}$, no general quantitative stability result.
2. Convex optimization viewpoint: $\boxed{\min_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle.}$
   $\rightarrow$ **Strong/Uniform convexity?** *Linear program.*
3. Elliptic PDE viewpoint: $\boxed{\det(\mathrm{D}^2\phi)\mu(\nabla\phi) = \rho.}$
   $\rightarrow$ **Uniform ellipticity?** *Degenerate elliptic equation.*

Using the unconstrained dual problem, quantitative stability guarantees can be obtained.

Introduction

- **Quantitative stability of optimal transport solutions w.r.t. problem data?**

  1. Beyond the case $\Omega = \mathbb{R}$, no general quantitative stability result.
  2. Convex optimization viewpoint: $\boxed{\min_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle.}$
     $\rightarrow$ **Strong/Uniform convexity?** *Linear program.*
  3. Elliptic PDE viewpoint: $\boxed{\det(\mathrm{D}^2 \phi)\mu(\nabla \phi) = \rho.}$
     $\rightarrow$ **Uniform ellipticity?** *Degenerate elliptic equation.*

Using the unconstrained dual problem, quantitative stability guarantees can be obtained.

Introduction

- **Quantitative stability of optimal transport solutions w.r.t. problem data?**

  1. Beyond the case $\Omega = \mathbb{R}$, no general quantitative stability result.
  2. Convex optimization viewpoint: $\boxed{\min_{\gamma \in \Gamma(\rho,\mu)} \langle d^2 | \gamma \rangle.}$
     $\rightarrow$ **Strong/Uniform convexity?** *Linear program.*
  3. Elliptic PDE viewpoint: $\boxed{\det(\mathrm{D}^2\phi)\mu(\nabla\phi) = \rho.}$
     $\rightarrow$ **Uniform ellipticity?** *Degenerate elliptic equation.*

Using the unconstrained dual problem, quantitative stability guarantees can be obtained.

Outline

Part I.
Strong convexity of the dual quadratic optimal transport problem.

Part II.
Consequences: quantitative stability estimates for optimal transport maps, Wasserstein barycenters and entropic optimal transport.

Outline

Part I.
Strong convexity of the dual quadratic optimal transport problem.

Part II.
Consequences: quantitative stability estimates for optimal transport maps, Wasserstein barycenters and entropic optimal transport.

Dual problem

$$\min_{\gamma\in\Gamma(\rho,\mu)}\int \|x-y\|^2\,\mathrm{d}\gamma(x,y) = \int \|x\|^2\,\mathrm{d}\rho(x) + \int \|y\|^2\,\mathrm{d}\mu(y) - 2\underbrace{\max_{\gamma\in\Gamma(\rho,\mu)}\int\langle x|y\rangle\mathrm{d}\gamma(x,y)}_{:=\mathcal{T}(\rho,\mu)}.$$

- Get rid of constraint:

$$\mathcal{T}(\rho,\mu) = \max_{\gamma\geq 0}\int\langle\cdot|\cdot\rangle\mathrm{d}\gamma + \min_{\phi,\psi}\int\phi\mathrm{d}(\rho-\gamma_1) + \int\psi\mathrm{d}(\mu-\gamma_2)$$

- Swap max and min:

$$\begin{aligned}\mathcal{T}(\rho,\mu) &= \min_{\phi,\psi}\int\phi\mathrm{d}\rho + \int\psi\mathrm{d}\mu + \max_{\gamma\geq 0}\int(\langle\cdot|\cdot\rangle - \phi\oplus\psi)\mathrm{d}\gamma \\ &= \min_{\phi,\psi|\langle\cdot|\cdot\rangle\leq\phi\oplus\psi}\int\phi\mathrm{d}\rho + \int\psi\mathrm{d}\mu \\ &= \min_{\psi}\int\psi^*\mathrm{d}\rho + \int\psi\mathrm{d}\mu, \quad \text{where} \quad \psi^*(x) = \sup_y\langle x|y\rangle - \psi(y).\end{aligned}$$

Remark: When ρ is a.c., $T = \nabla\psi^*$ where ψ is a minimizer.

Definition: $\mathcal{K}_\rho : \psi \mapsto \int \psi^*\mathrm{d}\rho.$

Dual problem

$$\min_{\gamma \in \Gamma(\rho,\mu)} \int \|x-y\|^2 \, \mathrm{d}\gamma(x,y) = \int \|x\|^2 \, \mathrm{d}\rho(x) + \int \|y\|^2 \, \mathrm{d}\mu(y) - 2 \underbrace{\max_{\gamma \in \Gamma(\rho,\mu)} \int \langle x|y \rangle \mathrm{d}\gamma(x,y)}_{:=\mathcal{T}(\rho,\mu)}.$$

- Get rid of constraint:

$$\mathcal{T}(\rho,\mu) = \max_{\gamma \geq 0} \int \langle \cdot|\cdot \rangle \mathrm{d}\gamma + \min_{\phi,\psi} \int \phi \mathrm{d}(\rho - \gamma_1) + \int \psi \mathrm{d}(\mu - \gamma_2)$$

- Swap max and min:

$$\begin{aligned}\mathcal{T}(\rho,\mu) &= \min_{\phi,\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \max_{\gamma \geq 0} \int (\langle \cdot|\cdot \rangle - \phi \oplus \psi) \mathrm{d}\gamma \\ &= \min_{\phi,\psi | \langle \cdot|\cdot \rangle \leq \phi \oplus \psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu \\ &= \min_{\psi} \int \psi^* \mathrm{d}\rho + \int \psi \mathrm{d}\mu, \quad \text{where} \quad \psi^*(x) = \sup_y \langle x|y \rangle - \psi(y).\end{aligned}$$

Remark: When ρ is a.c., $T = \nabla \psi^*$ where ψ is a minimizer.

Definition: $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho.$

Dual problem

$$\min_{\gamma\in\Gamma(\rho,\mu)} \int \|x-y\|^2 \mathrm{d}\gamma(x,y) = \int \|x\|^2 \mathrm{d}\rho(x) + \int \|y\|^2 \mathrm{d}\mu(y) - 2 \underbrace{\max_{\gamma\in\Gamma(\rho,\mu)} \int \langle x|y\rangle \mathrm{d}\gamma(x,y)}_{:=\mathcal{T}(\rho,\mu)}.$$

- Get rid of constraint:

$$\mathcal{T}(\rho,\mu) = \max_{\gamma\geq 0} \int \langle\cdot|\cdot\rangle \mathrm{d}\gamma + \min_{\phi,\psi} \int \phi \mathrm{d}(\rho-\gamma_1) + \int \psi \mathrm{d}(\mu-\gamma_2)$$

- Swap max and min:

$$\mathcal{T}(\rho,\mu) = \min_{\phi,\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \max_{\gamma\geq 0} \int (\langle\cdot|\cdot\rangle - \phi\oplus\psi)\mathrm{d}\gamma$$

$$= \min_{\phi,\psi|\langle\cdot|\cdot\rangle\leq\phi\oplus\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu$$

$$= \min_{\psi} \int \psi^* \mathrm{d}\rho + \int \psi \mathrm{d}\mu, \quad \text{where} \quad \psi^*(x) = \sup_y \langle x|y\rangle - \psi(y).$$

Remark: When ρ is a.c., $T = \nabla\psi^*$ where ψ is a minimizer.

Definition: $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho.$

Dual problem

$$\min_{\gamma\in\Gamma(\rho,\mu)} \int \|x-y\|^2 \,\mathrm{d}\gamma(x,y) = \int \|x\|^2 \,\mathrm{d}\rho(x) + \int \|y\|^2 \,\mathrm{d}\mu(y) - 2 \underbrace{\max_{\gamma\in\Gamma(\rho,\mu)} \int \langle x|y\rangle \mathrm{d}\gamma(x,y)}_{:=\mathcal{T}(\rho,\mu)}.$$

- Get rid of constraint:

$$\mathcal{T}(\rho,\mu) = \max_{\gamma\geq 0} \int \langle\cdot|\cdot\rangle \mathrm{d}\gamma + \min_{\phi,\psi} \int \phi \mathrm{d}(\rho-\gamma_1) + \int \psi \mathrm{d}(\mu-\gamma_2)$$

- Swap max and min:

$$\mathcal{T}(\rho,\mu) = \min_{\phi,\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \max_{\gamma\geq 0} \int (\langle\cdot|\cdot\rangle - \phi\oplus\psi)\mathrm{d}\gamma$$

$$= \min_{\phi,\psi | \langle\cdot|\cdot\rangle \leq \phi\oplus\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu$$

$$= \min_{\psi} \int \psi^* \mathrm{d}\rho + \int \psi \mathrm{d}\mu, \quad \text{where} \quad \psi^*(x) = \sup_y \langle x|y\rangle - \psi(y).$$

Remark: When ρ is a.c., $T = \nabla\psi^*$ where ψ is a minimizer.

Definition: $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho.$

Dual problem

$$\min_{\gamma\in\Gamma(\rho,\mu)} \int \|x-y\|^2 \,\mathrm{d}\gamma(x,y) = \int \|x\|^2 \,\mathrm{d}\rho(x) + \int \|y\|^2 \,\mathrm{d}\mu(y) - 2 \underbrace{\max_{\gamma\in\Gamma(\rho,\mu)} \int \langle x|y\rangle \mathrm{d}\gamma(x,y)}_{:=\mathcal{T}(\rho,\mu)}.$$

- Get rid of constraint:

$$\mathcal{T}(\rho,\mu) = \max_{\gamma\geq 0} \int \langle\cdot|\cdot\rangle \mathrm{d}\gamma + \min_{\phi,\psi} \int \phi \mathrm{d}(\rho-\gamma_1) + \int \psi \mathrm{d}(\mu-\gamma_2)$$

- Swap max and min:

$$\begin{aligned}\mathcal{T}(\rho,\mu) &= \min_{\phi,\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \max_{\gamma\geq 0} \int (\langle\cdot|\cdot\rangle - \phi\oplus\psi)\mathrm{d}\gamma \\ &= \min_{\phi,\psi|\langle\cdot|\cdot\rangle\leq\phi\oplus\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu \\ &= \min_{\psi} \int \psi^* \mathrm{d}\rho + \int \psi \mathrm{d}\mu, \quad \text{where} \quad \psi^*(x) = \sup_y \langle x|y\rangle - \psi(y).\end{aligned}$$

Remark: When ρ is a.c., $T = \nabla\psi^*$ where ψ is a minimizer.

Definition: $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho.$

Dual problem

$$\min_{\gamma\in\Gamma(\rho,\mu)} \int \|x-y\|^2 \,\mathrm{d}\gamma(x,y) = \int \|x\|^2 \,\mathrm{d}\rho(x) + \int \|y\|^2 \,\mathrm{d}\mu(y) - 2 \underbrace{\max_{\gamma\in\Gamma(\rho,\mu)} \int \langle x|y\rangle \mathrm{d}\gamma(x,y)}_{:=\mathcal{T}(\rho,\mu)}.$$

- Get rid of constraint:

$$\mathcal{T}(\rho,\mu) = \max_{\gamma\geq 0} \int \langle\cdot|\cdot\rangle \mathrm{d}\gamma + \min_{\phi,\psi} \int \phi \mathrm{d}(\rho-\gamma_1) + \int \psi \mathrm{d}(\mu-\gamma_2)$$

- Swap max and min:

$$\begin{aligned}\mathcal{T}(\rho,\mu) &= \min_{\phi,\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \max_{\gamma\geq 0} \int (\langle\cdot|\cdot\rangle - \phi\oplus\psi)\mathrm{d}\gamma \\ &= \min_{\phi,\psi \mid \langle\cdot|\cdot\rangle \leq \phi\oplus\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu \\ &= \min_{\psi} \int \psi^* \mathrm{d}\rho + \int \psi \mathrm{d}\mu, \quad \text{where} \quad \psi^*(x) = \sup_y \langle x|y\rangle - \psi(y).\end{aligned}$$

Remark: When ρ is a.c., $T = \nabla\psi^*$ where ψ is a minimizer.

Definition: $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho$.

Dual problem

$$\min_{\gamma\in\Gamma(\rho,\mu)} \int \|x-y\|^2 \,\mathrm{d}\gamma(x,y) = \int \|x\|^2 \,\mathrm{d}\rho(x) + \int \|y\|^2 \,\mathrm{d}\mu(y) - 2 \underbrace{\max_{\gamma\in\Gamma(\rho,\mu)} \int \langle x|y\rangle \mathrm{d}\gamma(x,y)}_{:=\mathcal{T}(\rho,\mu)}.$$

- Get rid of constraint:

$$\mathcal{T}(\rho,\mu) = \max_{\gamma\geq 0} \int \langle\cdot|\cdot\rangle \mathrm{d}\gamma + \min_{\phi,\psi} \int \phi \mathrm{d}(\rho-\gamma_1) + \int \psi \mathrm{d}(\mu-\gamma_2)$$

- Swap max and min:

$$\begin{aligned}\mathcal{T}(\rho,\mu) &= \min_{\phi,\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \max_{\gamma\geq 0} \int (\langle\cdot|\cdot\rangle - \phi\oplus\psi)\mathrm{d}\gamma \\ &= \min_{\phi,\psi \mid \langle\cdot|\cdot\rangle \leq \phi\oplus\psi} \int \phi \mathrm{d}\rho + \int \psi \mathrm{d}\mu \\ &= \min_{\psi} \int \psi^* \mathrm{d}\rho + \int \psi \mathrm{d}\mu, \quad \text{where} \quad \psi^*(x) = \sup_y \langle x|y\rangle - \psi(y).\end{aligned}$$

Remark: When ρ is a.c., $T = \nabla\psi^*$ where ψ is a minimizer.

Definition: $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho$.

Dual problem

$$\boxed{\min_{\psi:\mathbb{R}^d\to\bar{\mathbb{R}}} \mathcal{K}_\rho(\psi) + \int \psi \mathrm{d}\mu.} \tag{DP}$$

- $\mathcal{K}_\rho$ **is convex.** First order condition in (DP):

$$\boxed{\psi_\mu \text{ minimizer in (DP)} \iff 0 \in \partial\mathcal{K}_\rho(\psi_\mu) + \mu \iff \psi_\mu \in (\partial\mathcal{K}_\rho)^{-1}(-\mu).}$$

Stability estimate for $\mu \mapsto \psi_\mu$.

$\iff$

Strong/Uniform convexity estimate for $\psi \mapsto \mathcal{K}_\rho(\psi)$.

$F \in \mathcal{C}^2(\mathbb{R}^d)$ is strongly convex if there exists $\alpha > 0$ such that:

$$\forall x, y \in \mathbb{R}^d, t \in [0,1], \qquad \alpha\frac{t(1-t)}{2}\|x-y\|^2 \le (1-t)F(x) + tF(y) - F((1-t)x+ty),$$

$$\iff \forall x, y \in \mathbb{R}^d, \qquad \frac{\alpha}{2}\|x-y\|^2 \le F(y) - F(x) - \langle \nabla F(x) | y - x\rangle,$$

$$\iff \forall x \in \mathbb{R}^d, \qquad \alpha\mathrm{id} \preccurlyeq \nabla^2 F(x),$$

$$\iff \forall x, y \in \mathbb{R}^d, \qquad \|(\nabla F)^{-1}(y) - (\nabla F)^{-1}(x)\| \le \frac{1}{\alpha}\|y - x\|.$$

Dual problem

$$\boxed{\min_{\psi:\mathbb{R}^d\to\bar{\mathbb{R}}} \mathcal{K}_\rho(\psi) + \int \psi \mathrm{d}\mu.} \tag{DP}$$

- $\mathcal{K}_\rho$ **is convex.** First order condition in (DP):

$$\boxed{\psi_\mu \text{ minimizer in (DP)} \iff 0 \in \partial\mathcal{K}_\rho(\psi_\mu) + \mu \iff \psi_\mu \in (\partial\mathcal{K}_\rho)^{-1}(-\mu).}$$

Stability estimate for $\mu \mapsto \psi_\mu$.

$\iff$

Strong/Uniform convexity estimate for $\psi \mapsto \mathcal{K}_\rho(\psi)$.

$F \in \mathcal{C}^2(\mathbb{R}^d)$ is strongly convex if there exists $\alpha > 0$ such that:

$$\forall x, y \in \mathbb{R}^d, t \in [0,1], \qquad \alpha\frac{t(1-t)}{2}\|x-y\|^2 \leq (1-t)F(x) + tF(y) - F((1-t)x + ty),$$

$$\iff \forall x, y \in \mathbb{R}^d, \qquad \frac{\alpha}{2}\|x-y\|^2 \leq F(y) - F(x) - \langle \nabla F(x) | y - x\rangle,$$

$$\iff \forall x \in \mathbb{R}^d, \qquad \alpha \mathrm{id} \preccurlyeq \nabla^2 F(x),$$

$$\iff \forall x, y \in \mathbb{R}^d, \qquad \|(\nabla F)^{-1}(y) - (\nabla F)^{-1}(x)\| \leq \frac{1}{\alpha}\|y - x\|.$$

Dual problem

$$\min_{\psi:\mathbb{R}^d\to\bar{\mathbb{R}}} \mathcal{K}_\rho(\psi) + \int \psi \mathrm{d}\mu. \qquad \text{(DP)}$$

- $\mathcal{K}_\rho$ **is convex.** First order condition in (DP):

ψ_μ minimizer in (DP) $\iff 0 \in \partial\mathcal{K}_\rho(\psi_\mu) + \mu \iff \psi_\mu \in (\partial\mathcal{K}_\rho)^{-1}(-\mu).$

Stability estimate for $\mu \mapsto \psi_\mu$.

$\iff$

Strong/Uniform convexity estimate for $\psi \mapsto \mathcal{K}_\rho(\psi)$.

$F \in \mathcal{C}^2(\mathbb{R}^d)$ is strongly convex if there exists $\alpha > 0$ such that:

$$\forall x, y \in \mathbb{R}^d, t \in [0,1], \qquad \alpha\frac{t(1-t)}{2}\|x-y\|^2 \leq (1-t)F(x) + tF(y) - F((1-t)x+ty),$$

$$\iff \forall x, y \in \mathbb{R}^d, \qquad \frac{\alpha}{2}\|x-y\|^2 \leq F(y) - F(x) - \langle \nabla F(x) | y-x\rangle,$$

$$\iff \forall x \in \mathbb{R}^d, \qquad \alpha\mathrm{id} \preccurlyeq \nabla^2 F(x),$$

$$\iff \forall x, y \in \mathbb{R}^d, \qquad \|(\nabla F)^{-1}(y) - (\nabla F)^{-1}(x)\| \leq \frac{1}{\alpha}\|y-x\|.$$

Strong convexity of the dual problem

- **Strong convexity of $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho$?**
  1. *Strong convexity should work "up to additive constants":*

$$\boxed{\forall c \in \mathbb{R}, \quad \mathcal{K}_\rho(\psi + c) = \mathcal{K}_\rho(\psi) - c.}$$

  2. *Support of ρ should be connected:*

$$\nabla \psi^*_{\mu\#}\rho = \nabla \tilde{\psi}^*_{\mu\#}\rho = \mu.$$

$$\implies \psi_\mu, \tilde{\psi}_\mu \in \arg\min_\psi \mathcal{K}_\rho(\psi) + \langle \psi | \mu \rangle.$$

$$\implies \forall t \in [0,1],$$

$$\boxed{\mathcal{K}_\rho((1-t)\psi_\mu + t\tilde{\psi}_\mu) = (1-t)\mathcal{K}_\rho(\psi_\mu) + t\mathcal{K}_\rho(\tilde{\psi}_\mu).}$$

Assumption: Source ρ is **absolutely continuous** and satisfies a **Poincaré-Wirtinger inequality**: $\exists p \geq 1, C_{PW} \in (0, +\infty)$ s.t.

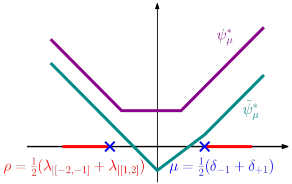
$$\forall f \in \mathcal{C}^1(\mathbb{R}^d), \quad \|f - \mathbb{E}_\rho f\|_{\mathrm{L}^p(\rho)} \leq C_{PW} \|\nabla f\|_{\mathrm{L}^p(\rho,\mathbb{R}^d)}.$$

Strong convexity of the dual problem

- **Strong convexity of $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho$?**
  1. *Strong convexity should work "up to additive constants":*

$$\boxed{\forall c \in \mathbb{R}, \quad \mathcal{K}_\rho(\psi + c) = \mathcal{K}_\rho(\psi) - c.}$$

  2. *Support of ρ should be connected:*



$$\nabla \psi_{\mu\#}^* \rho = \nabla \tilde{\psi}_{\mu\#}^* \rho = \mu.$$

$$\implies \psi_\mu, \tilde{\psi}_\mu \in \arg\min_\psi \mathcal{K}_\rho(\psi) + \langle \psi | \mu \rangle.$$

$$\implies \forall t \in [0,1],$$

$$\boxed{\mathcal{K}_\rho((1-t)\psi_\mu + t\tilde{\psi}_\mu) = (1-t)\mathcal{K}_\rho(\psi_\mu) + t\mathcal{K}_\rho(\tilde{\psi}_\mu).}$$

Assumption: Source ρ is **absolutely continuous** and satisfies a **Poincaré-Wirtinger inequality**: $\exists p \geq 1, C_{PW} \in (0, +\infty)$ s.t.

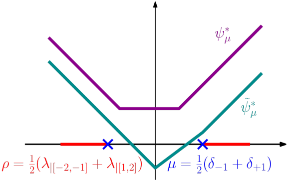
$$\forall f \in \mathcal{C}^1(\mathbb{R}^d), \quad \|f - \mathbb{E}_\rho f\|_{\mathrm{L}^p(\rho)} \leq C_{PW} \|\nabla f\|_{\mathrm{L}^p(\rho, \mathbb{R}^d)}.$$

Strong convexity of the dual problem

- **Strong convexity of $\mathcal{K}_\rho : \psi \mapsto \int \psi^* \mathrm{d}\rho$?**
  1. *Strong convexity should work "up to additive constants":*

$$\forall c \in \mathbb{R}, \quad \mathcal{K}_\rho(\psi + c) = \mathcal{K}_\rho(\psi) - c.$$

  2. *Support of ρ should be connected:*



$$\nabla\psi^*_{\mu\#}\rho = \nabla\tilde{\psi}^*_{\mu\#}\rho = \mu.$$

$$\implies \psi_\mu, \tilde{\psi}_\mu \in \arg\min_\psi \mathcal{K}_\rho(\psi) + \langle \psi | \mu \rangle.$$

$$\implies \forall t \in [0,1],$$

$$\mathcal{K}_\rho((1-t)\psi_\mu + t\tilde{\psi}_\mu) = (1-t)\mathcal{K}_\rho(\psi_\mu) + t\mathcal{K}_\rho(\tilde{\psi}_\mu).$$

Assumption: Source ρ is **absolutely continuous** and satisfies a **Poincaré-Wirtinger inequality**: $\exists p \geq 1, C_{PW} \in (0, +\infty)$ s.t.

$$\forall f \in \mathcal{C}^1(\mathbb{R}^d), \quad \|f - \mathbb{E}_\rho f\|_{\mathrm{L}^p(\rho)} \leq C_{PW} \|\nabla f\|_{\mathrm{L}^p(\rho, \mathbb{R}^d)}.$$

Strong convexity of the dual problem

- **"Subdifferential" of $\mathcal{K}_\rho$?** (Fenchel-Young) $\forall \psi, \tilde{\psi} : \mathbb{R}^d \to \bar{\mathbb{R}}$,

$$\langle \tilde{\psi} - \psi | -(\nabla \psi^*)_{\#}\rho \rangle \leq \mathcal{K}_\rho(\tilde{\psi}) - \mathcal{K}_\rho(\psi).$$

$\rightarrow$ Gap in this inequality?

- A known result:

Theorem (Ambrosio, Gigli, 2011):
Assume $\psi, \tilde{\psi} \in \mathcal{C}^1(\mathbb{R}^d)$, ψ is convex and $\tilde{\psi}$ **is α-strongly convex for some** $\alpha > 0$. Then,

$$\frac{\alpha}{2C_{PW}} \mathbb{V}\mathrm{ar}_\rho(\tilde{\psi}^* - \psi^*) \leq \frac{\alpha}{2} \left\| \nabla \tilde{\psi}^* - \nabla \psi^* \right\|^2_{\mathrm{L}^2(\rho, \mathbb{R}^d)} \leq \mathcal{K}_\rho(\tilde{\psi}) - \mathcal{K}_\rho(\psi) + \langle \tilde{\psi} - \psi | (\nabla \psi^*)_{\#}\rho \rangle.$$

Strong assumption: $\tilde{\psi}$ is α-strongly convex $\iff$ $\nabla \tilde{\psi}^*$ is $\frac{1}{\alpha}$-Lipschitz continuous!

$\rightarrow$ Not satisfied in general.

$\rightarrow$ Implies that $\nabla \tilde{\psi}^*_{\#}\rho$ has a connected support.

Strong convexity of the dual problem

- **"Subdifferential" of $\mathcal{K}_\rho$?** (Fenchel-Young) $\forall \psi, \tilde{\psi} : \mathbb{R}^d \to \bar{\mathbb{R}}$,

$$\langle \tilde{\psi} - \psi | -(\nabla \psi^*)_\# \rho \rangle \leq \mathcal{K}_\rho(\tilde{\psi}) - \mathcal{K}_\rho(\psi).$$

$\to$ Gap in this inequality?

- A known result:

Theorem (Ambrosio, Gigli, 2011):
Assume $\psi, \tilde{\psi} \in \mathcal{C}^1(\mathbb{R}^d)$, ψ is convex and $\tilde{\psi}$ **is α-strongly convex for some** $\alpha > 0$. Then,

$$\frac{\alpha}{2C_{PW}} \mathbb{V}\mathrm{ar}_\rho(\tilde{\psi}^* - \psi^*) \leq \frac{\alpha}{2} \left\| \nabla \tilde{\psi}^* - \nabla \psi^* \right\|^2_{\mathrm{L}^2(\rho, \mathbb{R}^d)} \leq \mathcal{K}_\rho(\tilde{\psi}) - \mathcal{K}_\rho(\psi) + \langle \tilde{\psi} - \psi | (\nabla \psi^*)_\# \rho \rangle.$$

Strong assumption: $\tilde{\psi}$ is α-strongly convex $\iff$ $\nabla \tilde{\psi}^*$ is $\frac{1}{\alpha}$-Lipschitz continuous!

$\to$ Not satisfied in general.
$\to$ Implies that $\nabla \tilde{\psi}^*_\# \rho$ has a connected support.

Strong convexity of the dual problem

Strong convexity estimate for the Kantorovich functional:

Theorem (D., Mérigot, 2021):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$.
- Let $\mathcal{Y} = B(0, R_\mathcal{Y}) \subset \mathbb{R}^d$ compact and let $\mu, \nu \in \mathcal{P}(\mathcal{Y})$.

Then for ψ_μ, ψ_ν Kantorovich potentials between ρ and μ, ν,

$$\boxed{C_{d,\mathcal{X},\rho,\mathcal{Y}} \mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu+\nu)}(\psi_\nu - \psi_\mu) \leq \mathcal{K}_\rho(\psi_\nu) - \mathcal{K}_\rho(\psi_\mu) + \langle \psi_\nu - \psi_\mu | \mu \rangle,}$$

where $C_{d,\mathcal{X},\rho,\mathcal{Y}} = \left(e(d+1)2^{d-1} R_\mathcal{Y} \mathrm{diam}(\mathcal{X}) \frac{M_\rho^2}{m_\rho^2} \right)^{-1}$.

Strong convexity of the dual problem

$$\boxed{\mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu+\nu)}(\psi_\nu - \psi_\mu) \lesssim \mathcal{K}_\rho(\psi_\nu) - \mathcal{K}_\rho(\psi_\mu) + \langle \psi_\nu - \psi_\mu | \mu \rangle.}$$

- Remarks:
  1. Reference $\frac{1}{2}(\mu+\nu)$? Fenchel-Young (in)equality:

Proposition (D., Mérigot, 2021): $\frac{1}{2}\mathbb{V}\mathrm{ar}_\rho(\psi_\mu^* - \psi_\nu^*) \leq \mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu+\nu)}(\psi_\nu - \psi_\mu)$.

  2. Optimal exponents.
  3. Similar result with **non-compact targets**: main assumption is boundedness of ψ_μ^*, ψ_ν^* on $\mathcal{X}$. Satisfied if μ, ν have bounded moment of order $p > d$ (Morrey's inequality).
  4. $\mathrm{spt}(\rho) = \mathcal{X}$ convex? Can be relaxed.

Corollary (Carlier, D., Mérigot, 2022): If $\mathcal{X}$ is a **connected finite union of convex sets** s.t. ρ satisfies a L^1 **Poincaré-Wirtinger inequality**, a similar estimate holds.

Strong convexity of the dual problem

$$\boxed{\mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu^0+\mu^1)}(\psi^1 - \psi^0) \lesssim \mathcal{K}_\rho(\psi^1) - \mathcal{K}_\rho(\psi^0) + \langle \psi^1 - \psi^0 | \mu^0 \rangle.}$$

- Elements of proof:
  1. Let $v := \psi^1 - \psi^0$. For $t \in [0,1], \psi^t := \psi^0 + tv$, $\phi^t := (\psi^t)^*$ and $\tilde{v}^t = v(\nabla \phi^t)$. Then *under regularity assumptions*:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_\rho(\psi^t) = -\mathbb{E}_\rho \tilde{v}^t \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t) = \mathbb{E}_\rho \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$$

  2. **Brascamp-Lieb inequality:** $\boxed{\mathbb{V}\mathrm{ar}_{e^{-\phi^t}}(\tilde{v}^t) \leq \mathbb{E}_{e^{-\phi^t}} \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$
  3. $m_\phi \leq \phi^t \leq M_\phi$ and $m_\rho \leq \rho \leq M_\rho \implies \boxed{\frac{e^{-M_\phi}}{M_\rho}\rho \leq e^{-\phi^t} \leq \frac{e^{-m_\phi}}{m_\rho}\rho.}$
  4. With $\mu^t = (\nabla\phi^t)_\# \rho$, $\boxed{\mathbb{V}\mathrm{ar}_{\mu^t}(v) \leq \frac{M_\rho}{m_\rho}\exp(M_\phi - m_\phi)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t).}$
  5. $\mu^t(y) = \rho(\nabla\psi^t(y)) \det(\mathrm{D}^2\psi^t(y))$ and log-concavity of $M \mapsto \det(M)$ over S.D.P. matrices $\implies \boxed{\mu^0 + \mu^1 \lesssim \mu^t.}$
  6. Conclusion: $\int_0^1 \ldots \mathrm{d}t$. General case by approximation.

Strong convexity of the dual problem

$$\boxed{\mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu^0+\mu^1)}(\psi^1 - \psi^0) \lesssim \mathcal{K}_\rho(\psi^1) - \mathcal{K}_\rho(\psi^0) + \langle \psi^1 - \psi^0 | \mu^0 \rangle.}$$

- Elements of proof:
  1. Let $v := \psi^1 - \psi^0$. For $t \in [0,1], \psi^t := \psi^0 + tv$, $\phi^t := (\psi^t)^*$ and $\tilde{v}^t = v(\nabla \phi^t)$. Then *under regularity assumptions*:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_\rho(\psi^t) = -\mathbb{E}_\rho \tilde{v}^t \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t) = \mathbb{E}_\rho \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$$

  2. **Brascamp-Lieb inequality:** $\boxed{\mathbb{V}\mathrm{ar}_{e^{-\phi^t}}(\tilde{v}^t) \leq \mathbb{E}_{e^{-\phi^t}} \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$
  3. $m_\phi \leq \phi^t \leq M_\phi$ and $m_\rho \leq \rho \leq M_\rho \implies \boxed{\frac{e^{-M_\phi}}{M_\rho}\rho \leq e^{-\phi^t} \leq \frac{e^{-m_\phi}}{m_\rho}\rho.}$
  4. With $\mu^t = (\nabla \phi^t)_\# \rho$, $\boxed{\mathbb{V}\mathrm{ar}_{\mu^t}(v) \leq \frac{M_\rho}{m_\rho}\exp(M_\phi - m_\phi)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t).}$
  5. $\mu^t(y) = \rho(\nabla \psi^t(y)) \det(\mathrm{D}^2\psi^t(y))$ and log-concavity of $M \mapsto \det(M)$ over S.D.P. matrices $\implies \boxed{\mu^0 + \mu^1 \lesssim \mu^t.}$
  6. Conclusion: $\int_0^1 ...\mathrm{d}t$. General case by approximation.

Strong convexity of the dual problem

$$\boxed{\mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu^0+\mu^1)}(\psi^1-\psi^0) \lesssim \mathcal{K}_\rho(\psi^1)-\mathcal{K}_\rho(\psi^0)+\langle\psi^1-\psi^0|\mu^0\rangle.}$$

- Elements of proof:
  1. Let $v := \psi^1 - \psi^0$. For $t \in [0,1], \psi^t := \psi^0 + tv$, $\phi^t := (\psi^t)^*$ and $\tilde{v}^t = v(\nabla\phi^t)$. Then *under regularity assumptions*:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_\rho(\psi^t) = -\mathbb{E}_\rho\tilde{v}^t \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t) = \mathbb{E}_\rho\langle\nabla\tilde{v}^t|(\mathrm{D}^2\phi^t)^{-1}\cdot\nabla\tilde{v}^t\rangle.}$$

  2. **Brascamp-Lieb inequality:** $\boxed{\mathbb{V}\mathrm{ar}_{e^{-\phi^t}}(\tilde{v}^t) \le \mathbb{E}_{e^{-\phi^t}}\langle\nabla\tilde{v}^t|(\mathrm{D}^2\phi^t)^{-1}\cdot\nabla\tilde{v}^t\rangle.}$
  3. $m_\phi \le \phi^t \le M_\phi$ and $m_\rho \le \rho \le M_\rho \implies \boxed{\frac{e^{-M_\phi}}{M_\rho}\rho \le e^{-\phi^t} \le \frac{e^{-m_\phi}}{m_\rho}\rho.}$
  4. With $\mu^t = (\nabla\phi^t)_\#\rho$, $\boxed{\mathbb{V}\mathrm{ar}_{\mu^t}(v) \le \frac{M_\rho}{m_\rho}\exp(M_\phi - m_\phi)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t).}$
  5. $\mu^t(y) = \rho(\nabla\psi^t(y))\det(\mathrm{D}^2\psi^t(y))$ and log-concavity of $M \mapsto \det(M)$ over S.D.P. matrices $\implies \boxed{\mu^0 + \mu^1 \lesssim \mu^t.}$
  6. Conclusion: $\int_0^1 ...\mathrm{d}t$. General case by approximation.

Strong convexity of the dual problem

$$\boxed{\mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu^0+\mu^1)}(\psi^1 - \psi^0) \lesssim \mathcal{K}_\rho(\psi^1) - \mathcal{K}_\rho(\psi^0) + \langle \psi^1 - \psi^0 | \mu^0 \rangle.}$$

- Elements of proof:
  1. Let $v := \psi^1 - \psi^0$. For $t \in [0,1], \psi^t := \psi^0 + tv$, $\phi^t := (\psi^t)^*$ and $\tilde{v}^t = v(\nabla\phi^t)$. Then *under regularity assumptions*:

     $$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_\rho(\psi^t) = -\mathbb{E}_\rho \tilde{v}^t \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t) = \mathbb{E}_\rho \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$$

  2. **Brascamp-Lieb inequality:** $\boxed{\mathbb{V}\mathrm{ar}_{e^{-\phi^t}}(\tilde{v}^t) \leq \mathbb{E}_{e^{-\phi^t}} \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$
  3. $m_\phi \leq \phi^t \leq M_\phi$ and $m_\rho \leq \rho \leq M_\rho \implies \boxed{\frac{e^{-M_\phi}}{M_\rho}\rho \leq e^{-\phi^t} \leq \frac{e^{-m_\phi}}{m_\rho}\rho.}$
  4. With $\mu^t = (\nabla\phi^t)_\#\rho$, $\boxed{\mathbb{V}\mathrm{ar}_{\mu^t}(v) \leq \frac{M_\rho}{m_\rho}\exp(M_\phi - m_\phi)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t).}$
  5. $\mu^t(y) = \rho(\nabla\psi^t(y))\det(\mathrm{D}^2\psi^t(y))$ and log-concavity of $M \mapsto \det(M)$ over S.D.P. matrices $\implies \boxed{\mu^0 + \mu^1 \lesssim \mu^t.}$
  6. Conclusion: $\int_0^1 ...\mathrm{d}t$. General case by approximation.

Strong convexity of the dual problem

$$\mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu^0+\mu^1)}(\psi^1 - \psi^0) \lesssim \mathcal{K}_\rho(\psi^1) - \mathcal{K}_\rho(\psi^0) + \langle \psi^1 - \psi^0 | \mu^0 \rangle.$$

- Elements of proof:
  1. Let $v := \psi^1 - \psi^0$. For $t \in [0,1], \psi^t := \psi^0 + tv$, $\phi^t := (\psi^t)^*$ and $\tilde{v}^t = v(\nabla \phi^t)$. Then *under regularity assumptions*:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_\rho(\psi^t) = -\mathbb{E}_\rho \tilde{v}^t \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t) = \mathbb{E}_\rho \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.$$

  2. **Brascamp-Lieb inequality:** $\mathbb{V}\mathrm{ar}_{e^{-\phi^t}}(\tilde{v}^t) \leq \mathbb{E}_{e^{-\phi^t}} \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.$
  3. $m_\phi \leq \phi^t \leq M_\phi$ and $m_\rho \leq \rho \leq M_\rho \implies \frac{e^{-M_\phi}}{M_\rho}\rho \leq e^{-\phi^t} \leq \frac{e^{-m_\phi}}{m_\rho}\rho.$
  4. With $\mu^t = (\nabla \phi^t)_{\#}\rho$, $\mathbb{V}\mathrm{ar}_{\mu^t}(v) \leq \frac{M_\rho}{m_\rho}\exp(M_\phi - m_\phi)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t).$
  5. $\mu^t(y) = \rho(\nabla \psi^t(y)) \det(\mathrm{D}^2\psi^t(y))$ and log-concavity of $M \mapsto \det(M)$ over S.D.P. matrices $\implies \mu^0 + \mu^1 \lesssim \mu^t.$
  6. Conclusion: $\int_0^1 ... \mathrm{d}t$. General case by approximation.

Strong convexity of the dual problem

$$\boxed{\mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu^0+\mu^1)}(\psi^1-\psi^0) \lesssim \mathcal{K}_\rho(\psi^1) - \mathcal{K}_\rho(\psi^0) + \langle \psi^1-\psi^0 | \mu^0 \rangle.}$$

- Elements of proof:
  1. Let $v := \psi^1 - \psi^0$. For $t \in [0,1], \psi^t := \psi^0 + tv$, $\phi^t := (\psi^t)^*$ and $\tilde{v}^t = v(\nabla \phi^t)$. Then *under regularity assumptions*:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_\rho(\psi^t) = -\mathbb{E}_\rho \tilde{v}^t \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t) = \mathbb{E}_\rho \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$$

  2. **Brascamp-Lieb inequality:** $\boxed{\mathbb{V}\mathrm{ar}_{e^{-\phi^t}}(\tilde{v}^t) \leq \mathbb{E}_{e^{-\phi^t}} \langle \nabla \tilde{v}^t | (\mathrm{D}^2\phi^t)^{-1} \cdot \nabla \tilde{v}^t \rangle.}$
  3. $m_\phi \leq \phi^t \leq M_\phi$ and $m_\rho \leq \rho \leq M_\rho \implies \boxed{\frac{e^{-M_\phi}}{M_\rho}\rho \leq e^{-\phi^t} \leq \frac{e^{-m_\phi}}{m_\rho}\rho.}$
  4. With $\mu^t = (\nabla \phi^t)_{\#}\rho$, $\boxed{\mathbb{V}\mathrm{ar}_{\mu^t}(v) \leq \frac{M_\rho}{m_\rho}\exp(M_\phi - m_\phi)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi^t).}$
  5. $\mu^t(y) = \rho(\nabla\psi^t(y))\det(\mathrm{D}^2\psi^t(y))$ and log-concavity of $M \mapsto \det(M)$ over S.D.P. matrices $\implies \boxed{\mu^0 + \mu^1 \lesssim \mu^t.}$
  6. Conclusion: $\int_0^1 ...\mathrm{d}t$. General case by approximation.

Outline

Part I.
Strong convexity of the dual quadratic optimal transport problem.

Part II.
Consequences: quantitative stability estimates for optimal transport maps, Wasserstein barycenters and entropic optimal transport.

Quantitative stability of optimal transport maps

- **Motivation: (Approximated) Wasserstein geometry.**

1. **Theorem** (Brenier, 1987): If $\rho \in \mathcal{P}_{2,a.c.}(\mathbb{R}^d)$, $\exists! \rho$-a.e. $T_\mu : \mathbb{R}^d \to \mathbb{R}^d$ s.t. $(T_\mu)_\# \rho = \mu$, and $T_\mu = \nabla \phi_\mu$ with ϕ_μ **convex**.

$$\implies \boxed{\mathrm{W}_2(\rho, \mu) = \| T_\mu - \mathrm{id} \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)} .}$$

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

	Riemannian geometry	Optimal transport
Point	$x \in M$	$\mu \in \mathcal{P}_2(\mathbb{R}^d)$
Geodesic distance	$d_g(x, y)$	$\mathrm{W}_2(\mu, \nu)$
Tangent space	$\mathcal{T}_\rho M$	$\mathcal{T}_\rho \mathcal{P}_2(\mathbb{R}^d) \subseteq \mathrm{L}^2(\rho, \mathbb{R}^d)$
Inverse exponential map	$\exp^{-1}(x) \in \mathcal{T}_\rho M$	$T_\mu - id \in \mathcal{T}_\rho \mathcal{P}_2(\mathbb{R}^d)$
Distance in tangent space	$\left\| \exp^{-1}(x) - \exp^{-1}(y) \right\|_{g(\rho)}$	$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}$

3. *Linearized Optimal Transport:* $\boxed{\mathrm{W}_{2,\rho}(\mu, \nu) := \| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)} .}$

Hilbertian distance used in image analysis (Wang et al., 2013) and other ML applications.

How much does $\mathrm{W}_{2,\rho}$ distort W_2?

(How stable is the mapping $\mu \mapsto T_\mu$?)

Quantitative stability of optimal transport maps

- **Motivation: (Approximated) Wasserstein geometry.**

1. **Theorem** (Brenier, 1987): If $\rho \in \mathcal{P}_{2,a.c.}(\mathbb{R}^d)$, $\exists! \rho$-a.e. $T_\mu : \mathbb{R}^d \to \mathbb{R}^d$ s.t. $(T_\mu)_\# \rho = \mu$, and $T_\mu = \nabla \phi_\mu$ with ϕ_μ **convex**.

$$\implies \boxed{\mathrm{W}_2(\rho, \mu) = \| T_\mu - \mathrm{id} \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}.}$$

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

	Riemannian geometry	Optimal transport
Point	$x \in M$	$\mu \in \mathcal{P}_2(\mathbb{R}^d)$
Geodesic distance	$d_g(x, y)$	$\mathrm{W}_2(\mu, \nu)$
Tangent space	$\mathcal{T}_\rho M$	$\mathcal{T}_\rho \mathcal{P}_2(\mathbb{R}^d) \subseteq \mathrm{L}^2(\rho, \mathbb{R}^d)$
Inverse exponential map	$\exp^{-1}(x) \in \mathcal{T}_\rho M$	$T_\mu - id \in \mathcal{T}_\rho \mathcal{P}_2(\mathbb{R}^d)$
Distance in tangent space	$\left\| \exp^{-1}(x) - \exp^{-1}(y) \right\|_{g(\rho)}$	$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}$

3. *Linearized Optimal Transport:* $\mathrm{W}_{2,\rho}(\mu, \nu) := \| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}.$

Hilbertian distance used in image analysis (Wang et al., 2013) and other ML applications.

How much does $\mathrm{W}_{2,\rho}$ distort W_2?

(How stable is the mapping $\mu \mapsto T_\mu$?)

Quantitative stability of optimal transport maps

- **Motivation: (Approximated) Wasserstein geometry.**

1. **Theorem** (Brenier, 1987): If $\rho \in \mathcal{P}_{2,a.c.}(\mathbb{R}^d)$, $\exists! \rho$-a.e. $T_\mu : \mathbb{R}^d \to \mathbb{R}^d$ s.t. $(T_\mu)_\# \rho = \mu$, and $T_\mu = \nabla \phi_\mu$ with ϕ_μ **convex**.

$$\Longrightarrow \boxed{\mathrm{W}_2(\rho, \mu) = \| T_\mu - \mathrm{id} \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}} .$$

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

	Riemannian geometry	Optimal transport
Point	$x \in M$	$\mu \in \mathcal{P}_2(\mathbb{R}^d)$
Geodesic distance	$d_g(x, y)$	$\mathrm{W}_2(\mu, \nu)$
Tangent space	$\mathcal{T}_\rho M$	$\mathcal{T}_\rho \mathcal{P}_2(\mathbb{R}^d) \subseteq \mathrm{L}^2(\rho, \mathbb{R}^d)$
Inverse exponential map	$\exp^{-1}(x) \in \mathcal{T}_\rho M$	$T_\mu - id \in \mathcal{T}_\rho \mathcal{P}_2(\mathbb{R}^d)$
Distance in tangent space	$\left\| \exp^{-1}(x) - \exp^{-1}(y) \right\|_{g(\rho)}$	$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}$

3. *Linearized Optimal Transport:* $\boxed{\mathrm{W}_{2,\rho}(\mu, \nu) := \| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)}}$.

Hilbertian distance used in image analysis (Wang et al., 2013) and other ML applications.

How much does $\mathrm{W}_{2,\rho}$ distort W_2?

(How stable is the mapping $\mu \mapsto T_\mu$?)

Quantitative stability of optimal transport maps

How much does $\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho,\mathbb{R}^d)}$ distort $\mathrm{W}_2(\mu,\nu)$?

- Elementary results: $\mu \mapsto T_\mu$ is **continuous** and **reverse-Lipschitz**.

$$\mathrm{W}_2(\mu,\nu) \leq \| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \,.$$

- Negative results: $\mu \mapsto T_\mu$ is **not better than $\frac{1}{2}$-Hölder**.

Theorem (Andoni, Naor, Neiman, 2018):
$(\mathcal{P}_2(\mathbb{R}^3), \mathrm{W}_2)$ does not admit a bi-Hölder embedding into any L^p space.

$\implies$ $\mu \mapsto T_\mu$ **is not bi-Hölder on the whole set** $\mathcal{P}_2(\mathbb{R}^d)$

Quantitative stability of optimal transport maps

How much does $\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho,\mathbb{R}^d)}$ distort $\mathrm{W}_2(\mu,\nu)$?

- Elementary results: $\mu \mapsto T_\mu$ is **continuous** and **reverse-Lipschitz**.

$$\mathrm{W}_2(\mu,\nu) \leq \| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} .$$

- Negative results: $\mu \mapsto T_\mu$ is **not better than $\frac{1}{2}$-Hölder**.

Theorem (Andoni, Naor, Neiman, 2018):
$(\mathcal{P}_2(\mathbb{R}^3), \mathrm{W}_2)$ does not admit a bi-Hölder embedding into any L^p space.

$\Longrightarrow$ $\mu \mapsto T_\mu$ **is not bi-Hölder on the whole set** $\mathcal{P}_2(\mathbb{R}^d)$

Quantitative stability of optimal transport maps

How much does $\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho,\mathbb{R}^d)}$ distort $\mathrm{W}_2(\mu,\nu)$?

- Known positive results:

Theorem (Ambrosio, Gigli, 2011): Let $\Omega \subset \mathbb{R}^d$ compact, $\rho \in \mathcal{P}_{a.c.}(\Omega)$ and $\mu, \nu \in \mathcal{P}(\Omega)$. Assume that T_μ is L-Lipschitz. Then,

$$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \leq 2(\mathrm{diam}(\Omega) L)^{1/2} \mathrm{W}_1(\mu,\nu)^{1/2}.$$

Theorem (Berman, 2020): Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$. Let $\mathcal{Y} \subset \mathbb{R}^d$ bounded connected open with Lipschitz boundary and $\mu, \nu \in \mathcal{P}(\mathcal{Y})$. Then,

$$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \leq C_{d,\rho,\mathcal{X},\mathcal{Y}} \mathrm{W}_1(\mu,\nu)^{\frac{1}{2^{(d-1)}(d+2)}}.$$

Quantitative stability of optimal transport maps

How much does $\|T_\mu - T_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)}$ distort $\mathrm{W}_2(\mu,\nu)$?

- Known positive results:

Theorem (Ambrosio, Gigli, 2011): Let $\Omega \subset \mathbb{R}^d$ compact, $\rho \in \mathcal{P}_{a.c.}(\Omega)$ and $\mu, \nu \in \mathcal{P}(\Omega)$. Assume that T_μ is L-Lipschitz. Then,

$$\|T_\mu - T_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \leq 2(\mathrm{diam}(\Omega)L)^{1/2}\mathrm{W}_1(\mu,\nu)^{1/2}.$$

Theorem (Berman, 2020): Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$. Let $\mathcal{Y} \subset \mathbb{R}^d$ bounded connected open with Lipschitz boundary and $\mu, \nu \in \mathcal{P}(\mathcal{Y})$. Then,

$$\|T_\mu - T_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \leq C_{d,\rho,\mathcal{X},\mathcal{Y}}\mathrm{W}_1(\mu,\nu)^{\frac{1}{2^{(d-1)}(d+2)}}.$$

Quantitative stability of optimal transport maps

Global & dimension-independent stability result (compact case):

Theorem (D., Mérigot, 2021):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$.
- Let $\mathcal{Y} = B(0, R_\mathcal{Y}) \subset \mathbb{R}^d$ bounded and $\mu, \nu \in \mathcal{P}(\mathcal{Y})$.

Then the OT maps T_μ, T_ν between ρ and μ, ν satisfy

$$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)} \leq C_{d,\rho,\mathcal{X},\mathcal{Y}} \mathrm{W}_1(\mu, \nu)^{1/6}.$$

- Remarks:
  1. Optimal Hölder exponent? $\frac{1}{6} < \frac{1}{2}$.
  2. The constant is explicit: $C_{d,\rho,\mathcal{X},\mathcal{Y}} \approx C_d \left(\frac{M_\rho}{m_\rho}\right)^3 \mathrm{diam}(\mathcal{X})^{d+1} R_\mathcal{Y}^2$.
  3. **Proof idea:** strong convexity of $\mathcal{K}_\rho$ and new Galgliardo-Nirenberg type inequality:

Proposition (D., Mérigot, 2021):
For $K \subset \mathbb{R}^d$ compact and $u, v : K \to \mathbb{R}$ Lipschitz convex,

$$\|\nabla u - \nabla v\|^2_{\mathrm{L}^2(K,\mathbb{R}^d)} \leq C_d \mathcal{H}^{d-1}(\partial K)^{\frac{2}{3}} (\mathrm{Lip}(u) + \mathrm{Lip}(v))^{\frac{4}{3}} \|u - v\|^{\frac{2}{3}}_{\mathrm{L}^2(K)}.$$

Quantitative stability of optimal transport maps

Global & dimension-independent stability result (compact case):

Theorem (D., Mérigot, 2021):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$.
- Let $\mathcal{Y} = B(0, R_{\mathcal{Y}}) \subset \mathbb{R}^d$ bounded and $\mu, \nu \in \mathcal{P}(\mathcal{Y})$.

Then the OT maps T_μ, T_ν between ρ and μ, ν satisfy

$$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)} \leq C_{d,\rho,\mathcal{X},\mathcal{Y}} \mathrm{W}_1(\mu, \nu)^{1/6}.$$

- Remarks:
  1. Optimal Hölder exponent? $\frac{1}{6} < \frac{1}{2}$.
  2. The constant is explicit: $C_{d,\rho,\mathcal{X},\mathcal{Y}} \approx C_d \left(\frac{M_\rho}{m_\rho} \right)^3 \mathrm{diam}(\mathcal{X})^{d+1} R_{\mathcal{Y}}^2$.
  3. **Proof idea:** strong convexity of $\mathcal{K}_\rho$ and new Galgliardo-Nirenberg type inequality:

Proposition (D., Mérigot, 2021):
For $K \subset \mathbb{R}^d$ compact and $u, v : K \to \mathbb{R}$ Lipschitz convex,

$$\|\nabla u - \nabla v\|^2_{\mathrm{L}^2(K,\mathbb{R}^d)} \leq C_d \mathcal{H}^{d-1}(\partial K)^{\frac{2}{3}} (\mathrm{Lip}(u) + \mathrm{Lip}(v))^{\frac{4}{3}} \|u - v\|^{\frac{2}{3}}_{\mathrm{L}^2(K)}.$$

Quantitative stability of optimal transport maps

Global stability result (non-compact case):

Theorem (Chazal, D., Mérigot, 2020/2021):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$.
- Let $p > d$ and $p \geq 4$ and let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with p-th moment upper bounded by $M_p < +\infty$.

Then the OT maps T_μ, T_ν between ρ and μ, ν satisfy

$$\| T_\mu - T_\nu \|_{\mathrm{L}^2(\rho, \mathbb{R}^d)} \leq C_{d,\rho,\mathcal{X},p,M_p} \mathrm{W}_1(\mu, \nu)^{\frac{p}{6p+16d}}.$$

- Remarks:
  1. Concerns all sub-Gaussian and sub-exponential probability measures.
  2. $\mu \mapsto T_\mu$ is a bi-Hölder embedding of $B_{\mathrm{W}_p}(\delta_0, M_p) \subset (\mathcal{P}_2(\mathbb{R}^d), \mathrm{W}_2)$ into $\mathrm{L}^2(\rho, \mathbb{R}^d)$.

Outline

Part I.
Strong convexity of the dual quadratic optimal transport problem.

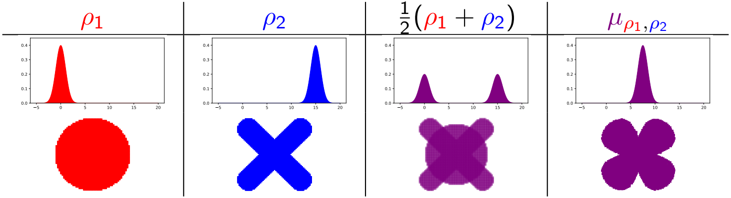
Part II.
Consequences: quantitative stability estimates for optimal transport maps, Wasserstein barycenters and entropic optimal transport.

Quantitative stability of Wasserstein barycenters

Definition: Let $\Omega \subset \mathbb{R}^d$ compact. *Wasserstein barycenter* of $\rho_1, \ldots, \rho_N \in \mathcal{P}(\Omega)$:

$$\mu_{\rho_1,\ldots,\rho_N} \in \arg \min_{\mu \in \mathcal{P}(\Omega)} \frac{1}{2N} \sum_{i=1}^{N} \mathrm{W}_2^2(\rho_i, \mu).$$

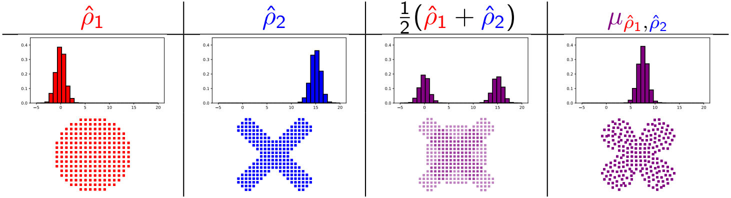
- Geometrically faithful "average" of probability measures:



Quantitative stability of Wasserstein barycenters

- Many applications, e.g. in
  1. Image processing (Rabin et al., 2011).
  2. Geometry processing (Solomon et al., 2015).
  3. Language processing (Colombo et al., 2021).

- ρ_1, ρ_2 often **not directly accessible**, but $\hat{\rho}_1$, $\hat{\rho}_2$ instead:



Can we bound $\mathrm{W}_2(\mu_{\hat{\rho}_1,\hat{\rho}_2}, \mu_{\rho_1,\rho_2})$ in terms of $\mathrm{W}_2(\hat{\rho}_1, \rho_1)$ and $\mathrm{W}_2(\hat{\rho}_2, \rho_2)$?

Quantitative stability of Wasserstein barycenters

- Known positive results:

Theorem (Le Gouic, Loubes, 2017): If $\forall i, \mathrm{W}_2(\rho_i^n, \rho_i) \xrightarrow[n\to\infty]{} 0$, then $(\mu_{\rho_1^n,\ldots,\rho_N^n})_n$ is pre-compact and any limit is a barycenter of $\rho_1, \ldots, \rho_N$.

Proposition: In dimension $d = 1$, $\mathrm{W}_2(\alpha, \beta) = \left\| F_\alpha^{-1} - F_\beta^{-1} \right\|_{\mathrm{L}^2([0,1])}$ so that:

$$\mathrm{W}_2(\mu_{\rho_1,\ldots,\rho_N}, \mu_{\tilde{\rho}_1,\ldots,\tilde{\rho}_N}) \leq \frac{1}{N} \sum_{i=1}^{N} \mathrm{W}_2(\rho_i, \tilde{\rho}_i).$$

Quantitative stability result in dimension $d \geq 2$?

Quantitative stability of Wasserstein barycenters

- Known positive results:

Theorem (Le Gouic, Loubes, 2017): If $\forall i, \mathrm{W}_2(\rho_i^n, \rho_i) \xrightarrow[n\to\infty]{} 0$, then $(\mu_{\rho_1^n,\dots,\rho_N^n})_n$ is pre-compact and any limit is a barycenter of $\rho_1, \dots, \rho_N$.

Proposition: In dimension $d = 1$, $\mathrm{W}_2(\alpha, \beta) = \left\| F_\alpha^{-1} - F_\beta^{-1} \right\|_{\mathrm{L}^2([0,1])}$ so that:

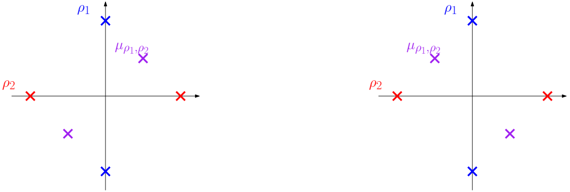
$$\mathrm{W}_2(\mu_{\rho_1,\dots,\rho_N}, \mu_{\tilde{\rho}_1,\dots,\tilde{\rho}_N}) \leq \frac{1}{N} \sum_{i=1}^{N} \mathrm{W}_2(\rho_i, \tilde{\rho}_i).$$

Quantitative stability result in dimension $d \geq 2$?

Quantitative stability of Wasserstein barycenters

- Negative results:

When $d \geq 2$, barycenter may not be unique:



Proposition (Agueh, Carlier, 2011): If one of the ρ_i's is absolutely continuous, the barycenter is unique.

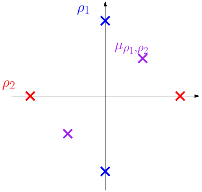
Even with an a.c. marginal, α-Hölder behaviour for any $\alpha \in (0,1)$ is possible:

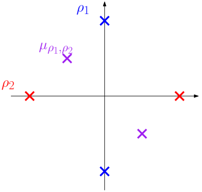
$W_2(\rho_2^0, \rho_2^\varepsilon) = \varepsilon$ while $W_2(\mu_{\rho_1,\rho_2^0}, \mu_{\rho_1,\rho_2^\varepsilon}) \sim \varepsilon^\alpha$.

Quantitative stability of Wasserstein barycenters

- Negative results:

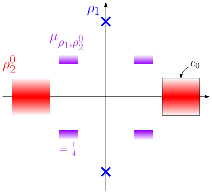
When $d \geq 2$, barycenter may not be unique:

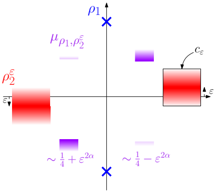




> **Proposition** (Agueh, Carlier, 2011): If one of the ρ_i's is absolutely continuous, the barycenter is unique.

Even with an a.c. marginal, α-Hölder behaviour for any $\alpha \in (0,1)$ is possible:





$$\mathrm{W}_2(\rho_2^0, \rho_2^\varepsilon) = \varepsilon \text{ while } \mathrm{W}_2(\mu_{\rho_1,\rho_2^0}, \mu_{\rho_1,\rho_2^\varepsilon}) \sim \varepsilon^\alpha.$$

Quantitative stability of Wasserstein barycenters

Hölder stability estimate for Wasserstein barycenters:

Theorem (Carlier, D., Mérigot, 2022):

- Let $\Omega = B(0, R_\Omega) \subset \mathbb{R}^d$ and $\rho_1, \ldots, \rho_N, \tilde{\rho}_1, \ldots, \tilde{\rho}_N \in \mathcal{P}(\Omega)$.
- Assume ρ_1 is a.c. with $0 < m_{\rho_1} \leq \rho_1 \leq M_{\rho_1}$ on $\mathcal{X}_1 = \mathrm{spt}(\rho_1)$ convex.

Then,
$$\boxed{\mathrm{W}_2(\mu_{\rho_1,\ldots,\rho_N}, \mu_{\tilde{\rho}_1,\ldots,\tilde{\rho}_N}) \leq C_{d,\Omega,\rho_1} \left(\sum_{i=1}^N \mathrm{W}_2(\rho_i, \tilde{\rho}_i) \right)^{1/6}.}$$

- Remarks:
  1. Optimal Hölder exponent?
  2. The constant is explicit: $C_{d,\Omega,\rho_1} \approx C_d (R_\Omega \mathrm{diam}(\mathcal{X}_1))^{d+1} \left(\frac{M_{\rho_1}}{m_{\rho_1}}\right)^{1/6}$.
  3. Main assumption on ρ_1: $\mathcal{K}_{\rho_1}$ should satisfy a strong convexity estimate.
  4. **Proof idea:** strong convexity of $\mathcal{K}_{\rho_1} \Rightarrow$ "strong convexity" of $\frac{1}{2}\mathrm{W}_2^2(\rho_1, \cdot)$:

Proposition (Carlier, D., Mérigot, 2022): Under the assumptions on ρ_1,

$$\forall \mu, \nu \in \mathcal{P}(\Omega), \quad \mathrm{W}_2^6(\mu,\nu) \lesssim \frac{1}{2}\mathrm{W}_2^2(\rho_1,\nu) - \frac{1}{2}\mathrm{W}_2^2(\rho_1,\mu) - \langle \frac{1}{2}\|\cdot\|^2 - \psi_{\rho_1\to\mu} | \nu - \mu \rangle.$$

Quantitative stability of Wasserstein barycenters

Hölder stability estimate for Wasserstein barycenters:

Theorem (Carlier, D., Mérigot, 2022):

- Let $\Omega = B(0, R_\Omega) \subset \mathbb{R}^d$ and $\rho_1, \ldots, \rho_N, \tilde{\rho}_1, \ldots, \tilde{\rho}_N \in \mathcal{P}(\Omega)$.
- Assume ρ_1 is a.c. with $0 < m_{\rho_1} \leq \rho_1 \leq M_{\rho_1}$ on $\mathcal{X}_1 = \mathrm{spt}(\rho_1)$ convex.

Then, $$\boxed{\mathrm{W}_2(\mu_{\rho_1,\ldots,\rho_N}, \mu_{\tilde{\rho}_1,\ldots,\tilde{\rho}_N}) \leq C_{d,\Omega,\rho_1} \left(\sum_{i=1}^N \mathrm{W}_2(\rho_i, \tilde{\rho}_i) \right)^{1/6}.}$$

- Remarks:
  1. Optimal Hölder exponent?
  2. The constant is explicit: $C_{d,\Omega,\rho_1} \approx C_d (R_\Omega \mathrm{diam}(\mathcal{X}_1))^{d+1} \left(\frac{M_{\rho_1}}{m_{\rho_1}}\right)^{1/6}$.
  3. Main assumption on ρ_1: $\mathcal{K}_{\rho_1}$ should satisfy a strong convexity estimate.
  4. **Proof idea:** strong convexity of $\mathcal{K}_{\rho_1} \Rightarrow$ "strong convexity" of $\frac{1}{2}\mathrm{W}_2^2(\rho_1, \cdot)$:

Proposition (Carlier, D., Mérigot, 2022): Under the assumptions on ρ_1,

$$\forall \mu, \nu \in \mathcal{P}(\Omega), \quad \mathrm{W}_2^6(\mu,\nu) \lesssim \frac{1}{2}\mathrm{W}_2^2(\rho_1,\nu) - \frac{1}{2}\mathrm{W}_2^2(\rho_1,\mu) - \langle \frac{1}{2}\|\cdot\|^2 - \psi_{\rho_1\to\mu} | \nu - \mu \rangle.$$

Quantitative stability of Wasserstein barycenters

- **Statistical consequence:**

Theorem (Fournier, Guillin, 2015):

- Let $\rho \in \mathcal{P}(\Omega)$ and $\hat{\rho}^n = \frac{1}{n}\sum_{j=1}^n \delta_{x_j}$ where $(x_j)_{1\leq j\leq n} \sim \rho^{\otimes n}$. Then:

$$\mathbb{E}\mathrm{W}_2^2(\hat{\rho}^n, \rho) \leq C_d R^2 \begin{cases} n^{-1/2} & \text{if } d < 4, \\ n^{-1/2}\log(n) & \text{if } d = 4, \\ n^{-2/d} & \text{else.} \end{cases}$$

Corollary (Carlier, D., Mérigot, 2022):

- Under the assumptions of the theorem, if $\forall i$, $\tilde{\rho}_i = \frac{1}{n}\sum_{j=1}^n \delta_{x_{i,j}}$ where $(x_{i,j})_{1\leq j\leq n} \sim \rho_i^{\otimes n}$, then

$$\mathbb{E}\mathrm{W}_2^2(\mu_{\rho_1,\ldots,\rho_N}, \mu_{\hat{\rho}_1^n,\ldots,\hat{\rho}_N^n}) \lesssim N^{1/3} \begin{cases} n^{-1/12} & \text{if } d < 4, \\ n^{-1/12}\log(n)^{1/6} & \text{if } d = 4, \\ n^{-1/(3d)} & \text{else.} \end{cases}$$

Outline

Part I.
Strong convexity of the dual quadratic optimal transport problem.

Part II.
Consequences: quantitative stability estimates for optimal transport maps, Wasserstein barycenters and entropic optimal transport.

Quantitative stability of entropic regularization

Entropic optimal transport: (Schrödinger, 1931; Léonard, 2014; Peyré and Cuturi, 2019)

- Primal and dual problems: for $\varepsilon > 0$,

$$\max_{\gamma \in \Gamma(\rho,\mu)} \int \langle x|y\rangle \mathrm{d}\gamma(x,y) - \varepsilon \mathrm{KL}(\gamma|\rho \otimes \mu) = \min_{\psi : \mathbb{R}^d \to \bar{\mathbb{R}}} \int \psi^{c,\varepsilon,\mu} \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \varepsilon,$$

where $\psi^{c,\varepsilon,\mu}(\cdot) = \varepsilon \log \int e^{\frac{\langle \cdot|y\rangle - \psi(y)}{\varepsilon}} \mathrm{d}\mu(y)$.

$$\mathcal{K}_\rho^{\varepsilon,\mu} : \psi \mapsto \int \psi^{c,\varepsilon,\mu} \mathrm{d}\rho.$$

Quantitative stability of entropic regularization

Entropic optimal transport: (Schrödinger, 1931; Léonard, 2014; Peyré and Cuturi, 2019)

- Primal and dual problems: for $\varepsilon > 0$,

$$\max_{\gamma \in \Gamma(\rho,\mu)} \int \langle x|y \rangle \mathrm{d}\gamma(x,y) - \varepsilon \mathrm{KL}(\gamma|\rho \otimes \mu) = \min_{\psi : \mathbb{R}^d \to \bar{\mathbb{R}}} \int \psi^{c,\varepsilon,\mu} \mathrm{d}\rho + \int \psi \mathrm{d}\mu + \varepsilon,$$

where $\psi^{c,\varepsilon,\mu}(\cdot) = \varepsilon \log \int e^{\frac{\langle \cdot | y \rangle - \psi(y)}{\varepsilon}} \mathrm{d}\mu(y)$.

$$\mathcal{K}_\rho^{\varepsilon,\mu} : \psi \mapsto \int \psi^{c,\varepsilon,\mu} \mathrm{d}\rho.$$

Strong convexity of the dual entropic problem

Strong convexity estimate for the entropic Kantorovich functional:

Theorem (D., 2022):

- Let $\mathcal{X}$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$.
- Let $\mathcal{Y} = B(0, R_\mathcal{Y})$ compact and let $\mu \in \mathcal{P}(\mathcal{Y})$.

Then for ψ_μ an entropic potential between ρ and μ and any $v : \mathbb{R}^d \to \mathbb{R}$,

$$\boxed{C_{\mathcal{X},\rho,\mathcal{Y},\varepsilon} \mathbb{V}\mathrm{ar}_\mu(v) \leq \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathcal{K}_\rho^{\varepsilon,\mu}(\psi_\mu + tv)\Big|_{t=0},}$$

where $C_{\mathcal{X},\rho,\mathcal{Y},\varepsilon} = \left(e^{R_\mathcal{Y} \mathrm{diam}(\mathcal{X})} \frac{M_\rho}{m_\rho} + \varepsilon\right)^{-1}$.

Remarks:

1. Allows to recover the *non-entropic* estimate when $\varepsilon \to 0$.
2. Constant improves when $\varepsilon \to 0$.

Strong convexity of the dual entropic problem

Strong convexity estimate for the entropic Kantorovich functional:

Theorem (D., 2022):

- Let $\mathcal{X}$ compact convex, $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$.
- Let $\mathcal{Y} = B(0, R_\mathcal{Y})$ compact and let $\mu \in \mathcal{P}(\mathcal{Y})$.

Then for ψ_μ an entropic potential between ρ and μ and any $v : \mathbb{R}^d \to \mathbb{R}$,

$$\boxed{C_{\mathcal{X},\rho,\mathcal{Y},\varepsilon} \mathbb{V}\mathrm{ar}_\mu(v) \leq \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathcal{K}_\rho^{\varepsilon,\mu}(\psi_\mu + tv)\Big|_{t=0}},$$

where $C_{\mathcal{X},\rho,\mathcal{Y},\varepsilon} = \left(e^{R_\mathcal{Y} \mathrm{diam}(\mathcal{X})} \frac{M_\rho}{m_\rho} + \varepsilon\right)^{-1}$.

Remarks:

1. Allows to recover the *non-entropic* estimate when $\varepsilon \to 0$.
2. Constant improves when $\varepsilon \to 0$.

Strong convexity of the dual entropic problem

$$\boxed{C_{\mathcal{X},\rho,\mathcal{Y},\varepsilon} \mathbb{V}\mathrm{ar}_\mu(v) \leq \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathcal{K}_\rho^{\varepsilon,\mu}(\psi_\mu + tv),}$$

- **Elements of proof:**
  1. Define $I : \psi \mapsto \log(\int_{\mathcal{X}} e^{-\psi^{c,\varepsilon,\mu}})$, and using the Prékopa-Leindler inequality, show

     $$\forall \varphi, \psi : \mathbb{R}^d \to \mathbb{R}, t \in [0,1], \quad I((1-t)\varphi + t\psi) \geq (1-t)I(\varphi) + tI(\psi).$$

  2. Noticing that I is $\mathcal{C}^2$, conclude with

     $$\frac{\mathrm{d}^2}{\mathrm{d}t^2} I(\psi_\mu + tv) = -\frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathcal{K}_{\tilde{\rho}}^{\varepsilon,\mu}(\psi_\mu + tv) + Var_{\tilde{\mu}}(v) \leq 0.$$

- **Remark:** with $I : \psi \mapsto \log \int e^{-\psi^*}$, this argument recovers Brascamp-Lieb from Prékopa-Leindler.

Strong convexity of the dual entropic problem

$$\boxed{C_{\mathcal{X},\rho,\mathcal{Y},\varepsilon}\mathbb{V}\mathrm{ar}_\mu(v) \leq \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho^{\varepsilon,\mu}(\psi_\mu + tv),}$$

- **Elements of proof:**
  1. Define $I : \psi \mapsto \log(\int_{\mathcal{X}} e^{-\psi^{c,\varepsilon,\mu}})$, and using the Prékopa-Leindler inequality, show

     $$\forall \varphi, \psi : \mathbb{R}^d \to \mathbb{R}, t \in [0,1], \quad I((1-t)\varphi + t\psi) \geq (1-t)I(\varphi) + tI(\psi).$$

  2. Noticing that I is $\mathcal{C}^2$, conclude with

     $$\frac{\mathrm{d}^2}{\mathrm{d}t^2}I(\psi_\mu + tv) = -\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_{\tilde{\rho}}^{\varepsilon,\mu}(\psi_\mu + tv) + Var_{\tilde{\mu}}(v) \leq 0.$$

- **Remark:** with $I : \psi \mapsto \log \int e^{-\psi^*}$, this argument recovers Brascamp-Lieb from Prékopa-Leindler.

Strong convexity of the dual entropic problem

$$\boxed{C_{\mathcal{X},\rho,\mathcal{Y},\varepsilon}\mathbb{V}\mathrm{ar}_\mu(v) \leq \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho^{\varepsilon,\mu}(\psi_\mu + tv),}$$

- **Elements of proof:**
  1. Define $I : \psi \mapsto \log(\int_{\mathcal{X}} e^{-\psi^{c,\varepsilon,\mu}})$, and using the Prékopa-Leindler inequality, show

     $$\forall \varphi, \psi : \mathbb{R}^d \to \mathbb{R}, t \in [0,1], \quad I((1-t)\varphi + t\psi) \geq (1-t)I(\varphi) + tI(\psi).$$

  2. Noticing that I is $\mathcal{C}^2$, conclude with

     $$\frac{\mathrm{d}^2}{\mathrm{d}t^2}I(\psi_\mu + tv) = -\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_{\tilde{\rho}}^{\varepsilon,\mu}(\psi_\mu + tv) + Var_{\tilde{\mu}}(v) \leq 0.$$

- **Remark:** with $I : \psi \mapsto \log \int e^{-\psi^*}$, this argument recovers Brascamp-Lieb from Prékopa-Leindler.

Quantitative stability of entropic regularization

- Denote $\psi^\varepsilon = \arg\min_\psi \mathcal{K}_\rho^{\varepsilon,\mu}(\psi) + \langle\psi|\mu\rangle + \varepsilon$. $\implies$ $\boxed{\nabla\mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon) + \mu = 0.}$
- Implicit function theorem $\implies$ $\boxed{\nabla^2\mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon)\dot{\psi^\varepsilon} + \frac{\partial}{\partial\varepsilon}(\nabla\mathcal{K}_\rho^{\varepsilon,\mu})(\psi^\varepsilon) = 0.}$

Theorem (D., 2022):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex. Let $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$ and assume ρ is α-Hölder continuous.
- Let $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.

Then $\varepsilon \mapsto \psi^\varepsilon$ is $\mathcal{C}^1$. For any $\varepsilon > 0$ and $\alpha' \in (0, \alpha)$,

$$\left\|\dot{\psi^\varepsilon}\right\|_2 \leq C_{\mathcal{X},\rho,\mathcal{Y},\mu} \min(\varepsilon^{\alpha'}, 1)$$

- **Remark:** $\varepsilon \mapsto \psi^\varepsilon$ is Lipschitz continuous $\to$ numerical optimal transport.

Corollary (D., 2022): Under the same assumptions, $\forall\varepsilon > 0$, $\forall\alpha' \in (0, \alpha)$,

$$\left\|\psi^\varepsilon - \psi^0\right\|_\infty \lesssim \varepsilon^{1+\alpha'},$$

$$\left|\mathrm{W}_{2,\varepsilon}^2(\rho,\mu) - \mathrm{W}_2^2(\rho,\mu) - C_{\rho,\mu}\varepsilon^2\right| \lesssim \varepsilon^{2+\alpha'},$$

where $\mathrm{W}_{2,\varepsilon}^2(\rho,\mu) = \int \|x-y\|^2 \,\mathrm{d}\gamma^\varepsilon(x,y)$. The last bound is tight.

Quantitative stability of entropic regularization

- Denote $\psi^\varepsilon = \arg\min_{\psi} \mathcal{K}_\rho^{\varepsilon,\mu}(\psi) + \langle \psi | \mu \rangle + \varepsilon. \implies \boxed{\nabla \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon) + \mu = 0.}$
- Implicit function theorem $\implies \boxed{\nabla^2 \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon)\dot{\psi^\varepsilon} + \frac{\partial}{\partial \varepsilon}(\nabla \mathcal{K}_\rho^{\varepsilon,\mu})(\psi^\varepsilon) = 0.}$

Theorem (D., 2022):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex. Let $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$ and assume ρ is α-Hölder continuous.
- Let $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.

Then $\varepsilon \mapsto \psi^\varepsilon$ is $\mathcal{C}^1$. For any $\varepsilon > 0$ and $\alpha' \in (0, \alpha)$,

$$\left\| \dot{\psi^\varepsilon} \right\|_2 \leq C_{\mathcal{X},\rho,\mathcal{Y},\mu} \min(\varepsilon^{\alpha'}, 1)$$

- **Remark:** $\varepsilon \mapsto \psi^\varepsilon$ is Lipschitz continuous $\to$ numerical optimal transport.

Corollary (D., 2022): Under the same assumptions, $\forall \varepsilon > 0$, $\forall \alpha' \in (0, \alpha)$,

$$\left\| \psi^\varepsilon - \psi^0 \right\|_\infty \lesssim \varepsilon^{1+\alpha'},$$

$$\left| \mathrm{W}_{2,\varepsilon}^2(\rho,\mu) - \mathrm{W}_2^2(\rho,\mu) - C_{\rho,\mu}\varepsilon^2 \right| \lesssim \varepsilon^{2+\alpha'},$$

where $\mathrm{W}_{2,\varepsilon}^2(\rho,\mu) = \int \|x-y\|^2 \, \mathrm{d}\gamma^\varepsilon(x,y)$. The last bound is tight.

Quantitative stability of entropic regularization

- Denote $\psi^\varepsilon = \arg\min_\psi \mathcal{K}_\rho^{\varepsilon,\mu}(\psi) + \langle \psi | \mu \rangle + \varepsilon$. $\implies$ $\boxed{\nabla \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon) + \mu = 0.}$
- Implicit function theorem $\implies$ $\boxed{\nabla^2 \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon)\dot{\psi^\varepsilon} + \frac{\partial}{\partial \varepsilon}(\nabla \mathcal{K}_\rho^{\varepsilon,\mu})(\psi^\varepsilon) = 0.}$

Theorem (D., 2022):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex. Let $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$ and assume ρ is α-Hölder continuous.
- Let $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.

Then $\varepsilon \mapsto \psi^\varepsilon$ is $\mathcal{C}^1$. For any $\varepsilon > 0$ and $\alpha' \in (0, \alpha)$,

$$\left\| \dot{\psi^\varepsilon} \right\|_2 \leq C_{\mathcal{X},\rho,\mathcal{Y},\mu} \min(\varepsilon^{\alpha'}, 1)$$

- **Remark:** $\varepsilon \mapsto \psi^\varepsilon$ is Lipschitz continuous $\to$ numerical optimal transport.

Corollary (D., 2022): Under the same assumptions, $\forall \varepsilon > 0$, $\forall \alpha' \in (0, \alpha)$,

$$\left\| \psi^\varepsilon - \psi^0 \right\|_\infty \lesssim \varepsilon^{1+\alpha'},$$

$$\left| \mathrm{W}_{2,\varepsilon}^2(\rho,\mu) - \mathrm{W}_2^2(\rho,\mu) - C_{\rho,\mu}\varepsilon^2 \right| \lesssim \varepsilon^{2+\alpha'},$$

where $\mathrm{W}_{2,\varepsilon}^2(\rho,\mu) = \int \|x-y\|^2 \, \mathrm{d}\gamma^\varepsilon(x,y)$. The last bound is tight.

Quantitative stability of entropic regularization

- Denote $\psi^\varepsilon = \arg\min_\psi \mathcal{K}_\rho^{\varepsilon,\mu}(\psi) + \langle \psi | \mu \rangle + \varepsilon. \implies \boxed{\nabla \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon) + \mu = 0.}$
- Implicit function theorem $\implies \boxed{\nabla^2 \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon)\dot{\psi^\varepsilon} + \frac{\partial}{\partial \varepsilon}(\nabla \mathcal{K}_\rho^{\varepsilon,\mu})(\psi^\varepsilon) = 0.}$

Theorem (D., 2022):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex. Let $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$ and assume ρ is α-Hölder continuous.
- Let $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.

Then $\varepsilon \mapsto \psi^\varepsilon$ is $\mathcal{C}^1$. For any $\varepsilon > 0$ and $\alpha' \in (0, \alpha)$,

$$\left\| \dot{\psi^\varepsilon} \right\|_2 \leq C_{\mathcal{X},\rho,\mathcal{Y},\mu} \min(\varepsilon^{\alpha'}, 1)$$

- **Remark:** $\varepsilon \mapsto \psi^\varepsilon$ is Lipschitz continuous $\to$ numerical optimal transport.

Corollary (D., 2022): Under the same assumptions, $\forall \varepsilon > 0$, $\forall \alpha' \in (0, \alpha)$,

$$\left\| \psi^\varepsilon - \psi^0 \right\|_\infty \lesssim \varepsilon^{1+\alpha'},$$

$$\left| \mathrm{W}_{2,\varepsilon}^2(\rho,\mu) - \mathrm{W}_2^2(\rho,\mu) - C_{\rho,\mu}\varepsilon^2 \right| \lesssim \varepsilon^{2+\alpha'},$$

where $\mathrm{W}_{2,\varepsilon}^2(\rho,\mu) = \int \|x-y\|^2 \mathrm{d}\gamma^\varepsilon(x,y)$. The last bound is tight.

Quantitative stability of entropic regularization

- Denote $\psi^\varepsilon = \arg\min_\psi \mathcal{K}_\rho^{\varepsilon,\mu}(\psi) + \langle \psi | \mu \rangle + \varepsilon.$ $\implies$ $\boxed{\nabla \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon) + \mu = 0.}$
- Implicit function theorem $\implies$ $\boxed{\nabla^2 \mathcal{K}_\rho^{\varepsilon,\mu}(\psi^\varepsilon)\dot{\psi}^\varepsilon + \frac{\partial}{\partial \varepsilon}(\nabla \mathcal{K}_\rho^{\varepsilon,\mu})(\psi^\varepsilon) = 0.}$

Theorem (D., 2022):

- Let $\mathcal{X} \subset \mathbb{R}^d$ compact convex. Let $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ with $0 < m_\rho \leq \rho \leq M_\rho$ and assume ρ is α-Hölder continuous.
- Let $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.

Then $\varepsilon \mapsto \psi^\varepsilon$ is $\mathcal{C}^1$. For any $\varepsilon > 0$ and $\alpha' \in (0, \alpha)$,

$$\left\| \dot{\psi}^\varepsilon \right\|_2 \leq C_{\mathcal{X},\rho,\mathcal{Y},\mu} \min(\varepsilon^{\alpha'}, 1)$$

- **Remark:** $\varepsilon \mapsto \psi^\varepsilon$ is Lipschitz continuous $\to$ numerical optimal transport.

Corollary (D., 2022): Under the same assumptions, $\forall \varepsilon > 0$, $\forall \alpha' \in (0, \alpha)$,

$$\left\| \psi^\varepsilon - \psi^0 \right\|_\infty \lesssim \varepsilon^{1+\alpha'},$$

$$\left| \mathrm{W}_{2,\varepsilon}^2(\rho,\mu) - \mathrm{W}_2^2(\rho,\mu) - C_{\rho,\mu}\varepsilon^2 \right| \lesssim \varepsilon^{2+\alpha'},$$

where $\mathrm{W}_{2,\varepsilon}^2(\rho,\mu) = \int \|x - y\|^2 \, \mathrm{d}\gamma^\varepsilon(x,y)$. The last bound is tight.

Conclusion

- **Main ideas**:
  1. **Dual quadratic OT problem** may enjoy form of **strong/uniform convexity**.
  2. Strong convexity deduced from **Brascamp-Lieb/Prékopa-Leindler** inequalities $\rightarrow$ reinforces the link between OT and these functional inequalities.
  3. Allows to prove **stability results in OT**:
     - Guarantees for statistical/numerical approximation of OT maps and Wasserstein barycenters.
     - Partial embedding of $(\mathcal{P}_2(\mathbb{R}^d), \mathrm{W}_2)$ into $\mathrm{L}^2(\rho, \mathbb{R}^d)$.
     - Quantitative control of entropic regularization.

- **Open questions:**
  1. **Generalizations of strong convexity results:** source measure, ground cost?
  2. **Numerical consequences of strong convexity result:** Newton methods, interior point methods, etc.

Thank you for your attention!

Appendix

Quantitative stability of optimal transport maps

$$\boxed{\|T_\mu - T_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \leq C_{d,\rho,\mathcal{X},\mathcal{Y}} \mathrm{W}_1(\mu,\nu)^{1/6}.}$$

- **Proof idea:** strong convexity of $\mathcal{K}_\rho$ and new Galgliardo-Nirenberg type inequality:

Proposition (D., Mérigot, 2021): For $K \subset \mathbb{R}^d$ compact and $u, v : K \to \mathbb{R}$ Lipschitz convex,

$$\|\nabla u - \nabla v\|^2_{\mathrm{L}^2(K,\mathbb{R}^d)} \leq C_d \mathcal{H}^{d-1}(\partial K)^{\frac{2}{3}} (\mathrm{Lip}(u) + \mathrm{Lip}(v))^{\frac{4}{3}} \|u - v\|^{\frac{2}{3}}_{\mathrm{L}^2(K)}.$$

Let ϕ_μ, ϕ_ν be convex Brenier potentials between ρ and μ, ν:

$$\begin{aligned}
\|T_\mu - T_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} &= \|\nabla\phi_\mu - \nabla\phi_\nu\|_{\mathrm{L}^2(\rho,\mathbb{R}^d)} \\
&\lesssim \mathbb{V}\mathrm{ar}_\rho(\phi_\mu - \phi_\nu)^{1/6} \\
&\leq \mathbb{V}\mathrm{ar}_{\frac{1}{2}(\mu+\nu)}(\phi^*_\mu - \phi^*_\nu)^{1/6} \\
&\lesssim (\mathcal{K}_\rho(\psi_\nu) - \mathcal{K}_\rho(\psi_\mu) + \langle \psi_\nu - \psi_\mu | \mu\rangle)^{1/6} \\
&\lesssim \mathrm{W}_1(\mu,\nu)^{1/6}.
\end{aligned}$$

Quantitative stability of Wasserstein barycenters

▶ **General result: infinite number of marginals**

Definition: Let $\Omega \subset \mathbb{R}^d$ compact. *Wasserstein barycenter* of $\mathbb{P} \in \mathcal{P}(\mathcal{P}(\Omega))$:

$$\mu_{\mathbb{P}} \in \arg\min_{\mu \in \mathcal{P}(\Omega)} \frac{1}{2} \int_{\mathcal{P}(\Omega)} \mathrm{W}_2^2(\rho, \mu) \mathrm{d}\mathbb{P}(\rho).$$

Theorem (Carlier, D., Mérigot, 2022):

▶ *Let* $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\mathcal{P}(\Omega))$ *and* $\exists S_{\mathbb{P}} \subset \mathcal{P}(\Omega)$ *s.t.* $\mathbb{P}(S_{\mathbb{P}}) = \alpha_{\mathbb{P}} > 0$ *and* $\forall \rho \in S_{\mathbb{P}}$,

1. ρ *is a.c.*
2. $0 < m \leq \rho \leq M < +\infty$ *on its support.*
3. $\mathcal{H}^{d-1}(\partial \mathrm{spt}(\rho)) \leq \mathrm{per} < +\infty$.
4. $\forall \psi, \tilde{\psi} \in \mathcal{C}(\Omega)$,

$$c \mathbb{V}\mathrm{ar}_\rho(\tilde{\psi}^* - \psi^*) \leq \mathcal{K}_\rho(\tilde{\psi}) - \mathcal{K}_\rho(\psi) - \langle \tilde{\psi} - \psi | \nabla \mathcal{K}_\rho(\psi) \rangle.$$

Then,

$$\mathrm{W}_2(\mu_{\mathbb{P}}, \mu_{\mathbb{Q}}) \lesssim \frac{1}{\alpha_{\mathbb{P}}^{1/4}} \mathcal{W}_1(\mathbb{P}, \mathbb{Q})^{1/6}.$$

$$\forall \mathbb{P}, \mathbb{Q} \in \mathcal{P}(\mathcal{P}(\Omega)), \quad \mathcal{W}_1(\mathbb{P}, \mathbb{Q}) := \min_{\gamma \in \Pi(\mathbb{P}, \mathbb{Q})} \int_{\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)} \mathrm{W}_2(\rho, \tilde{\rho}) \mathrm{d}\gamma(\rho, \tilde{\rho}).$$

Quantitative stability of entropic regularization

- Let $c(x,y) = \|x-y\|^2$.

	Classical OT	*Entropic* OT
Problem	$(\mathrm{P}_0)_{\rho,\mu} : \min_{\gamma\in\Gamma(\rho,\mu)} \langle c\|\gamma\rangle$	$(\mathrm{P}_\varepsilon)_{\rho,\mu} : \min_{\gamma\in\Gamma(\rho,\mu)} \langle c\|\gamma\rangle + \varepsilon \mathrm{KL}(\gamma\|\rho\otimes\mu)$
Computational complexity (n support points)	$\tilde{O}(n^3)$	$\tilde{O}(n^2/\varepsilon^2)$
Sample complexity $\mathbb{E}\left\|(\mathrm{P}.)_{\hat\rho^n,\hat\mu^n} - (\mathrm{P}.)_{\rho,\mu}\right\|$ ($\hat\rho^n = \frac{1}{n}\Sigma_i \delta_{x_i}, x_i \sim \rho$, $\hat\mu^n = \frac{1}{n}\Sigma_j \delta_{y_j}, y_j \sim \mu$)	$O(n^{-1/d})$	$\lesssim \frac{e^{1/\varepsilon}}{\varepsilon^{d/2}} n^{-1/2}$
Minimizer	γ^0	γ^ε
Geometry?	$\mathrm{W}_2(\rho,\mu) := \sqrt{\langle c\|\gamma^0\rangle}$ is a distance	$\mathrm{W}_{2,\varepsilon}(\rho,\mu) := \sqrt{\langle c\|\gamma^\varepsilon\rangle}$ is not a distance

Quantitative stability of entropic regularization

- **Semi-discrete setting:**
  - Let $\mathcal{X} \subset \mathbb{R}^d$ compact and $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ absolutely continuous.
  - Let $\mathcal{Y} = \{y_1, \ldots, y_N\} \subset \mathbb{R}^d$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i} \in \mathcal{P}(\mathcal{Y})$.

  $\rightarrow$ Natural framework in **statistics** and **numerical analysis**.

- A 2-dimensional example:

$\rho = \mathbb{1}_{[a,b]\times[c,d]}$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.

($\varepsilon = 0$)

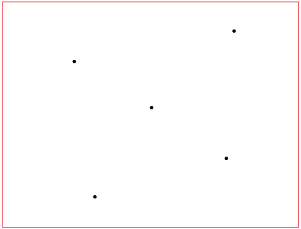
$T_{\rho\to\mu}$ is piece-wise constant.

Quantitative stability of entropic regularization

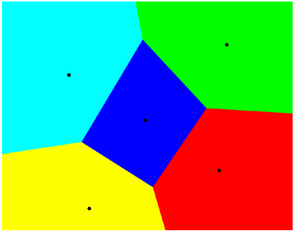
- **Semi-discrete setting:**
  - Let $\mathcal{X} \subset \mathbb{R}^d$ compact and $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ absolutely continuous.
  - Let $\mathcal{Y} = \{y_1, \ldots, y_N\} \subset \mathbb{R}^d$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i} \in \mathcal{P}(\mathcal{Y})$.

$\rightarrow$ Natural framework in **statistics** and **numerical analysis**.

- A 2-dimensional example:



$\rho = \mathbb{1}_{[a,b]\times[c,d]}$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.



$(\varepsilon = 0)$

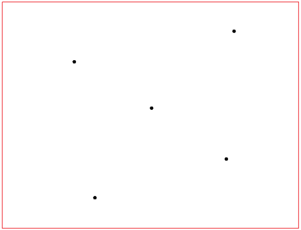
$T_{\rho\rightarrow\mu}$ is piece-wise constant.

Quantitative stability of entropic regularization

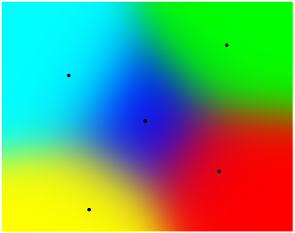
- **Semi-discrete setting:**
  - Let $\mathcal{X} \subset \mathbb{R}^d$ compact and $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ absolutely continuous.
  - Let $\mathcal{Y} = \{y_1, \ldots, y_N\} \subset \mathbb{R}^d$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i} \in \mathcal{P}(\mathcal{Y})$.

  $\rightarrow$ Natural framework in **statistics** and **numerical analysis**.

- A 2-dimensional example:



$\rho = \mathbb{1}_{[a,b]\times[c,d]}$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.



($\varepsilon = 7.5 \times 10^{-2}$)

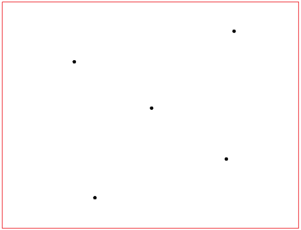
γ^ε has same support as $\rho \otimes \mu$.

Quantitative stability of entropic regularization

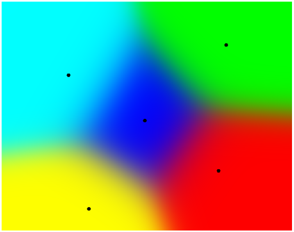
- **Semi-discrete setting:**
  - Let $\mathcal{X} \subset \mathbb{R}^d$ compact and $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ absolutely continuous.
  - Let $\mathcal{Y} = \{y_1, \dots, y_N\} \subset \mathbb{R}^d$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i} \in \mathcal{P}(\mathcal{Y})$.

  $\rightarrow$ Natural framework in **statistics** and **numerical analysis**.

- A 2-dimensional example:



$\rho = \mathbb{1}_{[a,b]\times[c,d]}$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.



$(\varepsilon = 5 \times 10^{-2})$

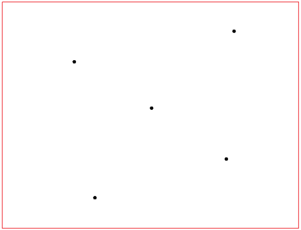
γ^ε has same support as $\rho \otimes \mu$.

Quantitative stability of entropic regularization

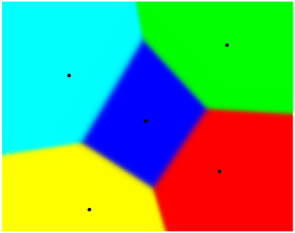
- **Semi-discrete setting:**
  - Let $\mathcal{X} \subset \mathbb{R}^d$ compact and $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ absolutely continuous.
  - Let $\mathcal{Y} = \{y_1, \dots, y_N\} \subset \mathbb{R}^d$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i} \in \mathcal{P}(\mathcal{Y})$.

$\rightarrow$ Natural framework in **statistics** and **numerical analysis**.

- A 2-dimensional example:



$\rho = \mathbb{1}_{[a,b]\times[c,d]}$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.



($\varepsilon = 7.5 \times 10^{-3}$)

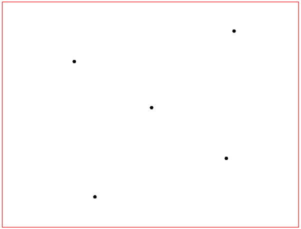
γ^ε has same support as $\rho \otimes \mu$.

Quantitative stability of entropic regularization

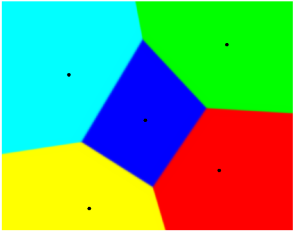
- **Semi-discrete setting:**
  - Let $\mathcal{X} \subset \mathbb{R}^d$ compact and $\rho \in \mathcal{P}_{a.c.}(\mathcal{X})$ absolutely continuous.
  - Let $\mathcal{Y} = \{y_1, \ldots, y_N\} \subset \mathbb{R}^d$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i} \in \mathcal{P}(\mathcal{Y})$.

$\rightarrow$ Natural framework in **statistics** and **numerical analysis**.

- A 2-dimensional example:



$\rho = \mathbb{1}_{[a,b]\times[c,d]}$ and $\mu = \sum_{i=1}^N \mu_i \delta_{y_i}$.



$(\varepsilon = 2.5 \times 10^{-3})$

γ^ε has same support as $\rho \otimes \mu$.