Convergence rate of entropy-regularized multi-marginal optimal transport costs

CANUM 2024

Paul Pegon

Joint work with L. Nenna

May 31, 2024

CEREMADE, Université Paris-Dauphine & MOKAPLAN, Inria Paris

Overview

1. Entropic Multi-Marginal Optimal Transport

2. The upper bound

3. The lower bound

Entropic Multi-Marginal Optimal Transport

Consider

- $m \geq 2$ probability measures μ_i compactly supported on $\mathscr{C}^2$ submanifolds $X_i \subseteq \mathbb{R}^N$ of dim d_i ;
- a cost function $c : \mathsf{X} \to \mathbb{R}_+$ (e.g. continuous or lsc) where $\mathsf{X} := \times_i^m X_i$;

Definition of the problem

Consider

- $m \geq 2$ probability measures μ_i compactly supported on $\mathscr{C}^2$ submanifolds $X_i \subseteq \mathbb{R}^N$ of dim d_i ;
- a cost function $c : \mathsf{X} \to \mathbb{R}_+$ (e.g. continuous or lsc) where $\mathsf{X} := \times_i^m X_i$;

Entropic Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_\varepsilon := \inf_{\gamma \in \Pi(\mu_1, \dots, \mu_m)} \left\{ \int_{\mathsf{X}} c(x_1, \dots, x_m) \, \mathrm{d}\gamma(x_1, \dots, x_m) + \varepsilon \mathrm{Ent}(\gamma \,|\, \otimes_{i=1}^m \mu_i) \right\},$$

where $\Pi(\mu_1, \dots, \mu_m)$ is the set of ***couplings*** $\gamma \in \mathscr{P}(\mathsf{X})$ having μ_i as marginals, $\mathrm{Ent}(\gamma \,|\, P)$ is the Boltzmann-Shannon entropy ($= \int \rho \log \rho \, \mathrm{d}P$ if $\gamma = \rho P$), and $\varepsilon > 0$ is a small noise parameter.

Consider

- $m \geq 2$ probability measures μ_i compactly supported on $\mathscr{C}^2$ submanifolds $X_i \subseteq \mathbb{R}^N$ of dim d_i ;
- a cost function $c : \mathsf{X} \to \mathbb{R}_+$ (e.g. continuous or lsc) where $\mathsf{X} := \times_i^m X_i$;

Entropic Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_\varepsilon := \inf_{\gamma \in \Pi(\mu_1,\ldots,\mu_m)} \left\{ \int_{\mathsf{X}} c(x_1,\ldots,x_m)\, \mathrm{d}\gamma(x_1,\ldots,x_m) + \varepsilon \mathrm{Ent}(\gamma \,|\, \otimes_{i=1}^m \mu_i) \right\},$$

where $\Pi(\mu_1,\ldots,\mu_m)$ is the set of ***couplings*** $\gamma \in \mathscr{P}(\mathsf{X})$ having μ_i as marginals, $\mathrm{Ent}(\gamma \,|\, P)$ is the Boltzmann-Shannon entropy ($= \int \rho \log \rho \,\mathrm{d}P$ if $\gamma = \rho P$), and $\varepsilon > 0$ is a small noise parameter.

- $m = 2$. ***Classical Entropic Optimal Transport.***

Consider

- $m \geq 2$ probability measures μ_i compactly supported on $\mathscr{C}^2$ submanifolds $X_i \subseteq \mathbb{R}^N$ of dim d_i ;
- a cost function $c : \mathsf{X} \to \mathbb{R}_+$ (e.g. continuous or lsc) where $\mathsf{X} := \times_i^m X_i$;

Entropic Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_\varepsilon := \inf_{\gamma \in \Pi(\mu_1,\dots,\mu_m)} \left\{ \int_{\mathsf{X}} c(x_1,\dots,x_m)\, \mathrm{d}\gamma(x_1,\dots,x_m) + \varepsilon \mathrm{Ent}(\gamma \mid \otimes_{i=1}^m \mu_i) \right\},$$

where $\Pi(\mu_1,\dots,\mu_m)$ is the set of ***couplings*** $\gamma \in \mathscr{P}(\mathsf{X})$ having μ_i as marginals, $\mathrm{Ent}(\gamma \mid P)$ is the Boltzmann-Shannon entropy ($= \int \rho \log \rho \, \mathrm{d}P$ if $\gamma = \rho P$), and $\varepsilon > 0$ is a small noise parameter.

- $m = 2$. ***Classical Entropic Optimal Transport.***
- $\varepsilon = 0$ and $m = 2$. ***Classical Optimal Transport problem.*** Convex problem, but may have several solutions γ, with or without finite entropy!

Definition of the problem

Consider

- $m \geq 2$ probability measures μ_i compactly supported on $\mathscr{C}^2$ submanifolds $X_i \subseteq \mathbb{R}^N$ of dim d_i ;
- a cost function $c : \mathsf{X} \to \mathbb{R}_+$ (e.g. continuous or lsc) where $\mathsf{X} := \times_i^m X_i$;

Entropic Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_\varepsilon := \inf_{\gamma \in \Pi(\mu_1,\dots,\mu_m)} \left\{ \int_{\mathsf{X}} c(x_1,\dots,x_m)\, \mathrm{d}\gamma(x_1,\dots,x_m) + \varepsilon \mathrm{Ent}(\gamma \,|\, \otimes_{i=1}^m \mu_i) \right\},$$

where $\Pi(\mu_1,\dots,\mu_m)$ is the set of ***couplings*** $\gamma \in \mathscr{P}(\mathsf{X})$ having μ_i as marginals, $\mathrm{Ent}(\gamma \,|\, P)$ is the Boltzmann-Shannon entropy ($= \int \rho \log \rho \, \mathrm{d}P$ if $\gamma = \rho P$), and $\varepsilon > 0$ is a small noise parameter.

- $m = 2$. ***Classical Entropic Optimal Transport.***
- $\varepsilon = 0$ and $m = 2$. ***Classical Optimal Transport problem.*** Convex problem, but may have several solutions γ, with or without finite entropy!
- $\varepsilon > 0$. Strictly convex cost $\implies$ unique solution γ_ε with finite entropy.

What are we interested in?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

What are we interested in?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

What are we interested in?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

First remark: depends heavily on $c = c(\boldsymbol{x})$ and the marginals μ_i's.

What are we interested in?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

First remark: depends heavily on $c = c(\boldsymbol{x})$ and the marginals μ_i's.

What is known?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

First remark: depends heavily on $c = c(\mathbf{x})$ and the marginals μ_i's.
What is known? Mostly for the $m = 2$ case.

What are we interested in?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

First remark: depends heavily on $c = c(\mathbf{x})$ and the marginals μ_i's.
What is known? Mostly for the $m = 2$ case.

- Convergence rate in the discrete case (Cominetti, San Martin ; 1994) and semi-discrete case (Altschuler, Niles-Weed, Stromme ; 2021) (Delalande ; 2021);

What are we interested in?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

First remark: depends heavily on $c = c(\mathbf{x})$ and the marginals μ_i's.
What is known? Mostly for the $m = 2$ case.

- Convergence rate in the discrete case (Cominetti, San Martin ; 1994) and semi-discrete case (Altschuler, Niles-Weed, Stromme ; 2021) (Delalande ; 2021);
- Convergence rate in the continuous case under strong regularity assumptions (Pal, 2019), second-order expansion for dynamical quadratic OT (Conforti-Tamanini, 2019)

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

First remark: depends heavily on $c = c(\mathbf{x})$ and the marginals μ_i's.
What is known? Mostly for the $m = 2$ case.

- Convergence rate in the discrete case (Cominetti, San Martin ; 1994) and semi-discrete case (Altschuler, Niles-Weed, Stromme ; 2021) (Delalande ; 2021);
- Convergence rate in the continuous case under strong regularity assumptions (Pal, 2019), second-order expansion for dynamical quadratic OT (Conforti-Tamanini, 2019)
- Convergence rate for 2−marginal and a general class of $\mathscr{C}^2$ non-degenerate costs (Carlier, Pegon, and Tamanini 2022), more precise estimates by (Malamut, Sylvestre ; 2023)

What are we interested in?

Asymptotics for $\varepsilon \to 0$ of

- the cost MOT_ε
- the optimal entropic plan γ_ε and optimal Schrödinger potentials $(\phi_\varepsilon, \psi_\varepsilon)$

First remark: depends heavily on $c = c(\mathbf{x})$ and the marginals μ_i's.
What is known? Mostly for the $m = 2$ case.

- Convergence rate in the discrete case (Cominetti, San Martin ; 1994) and semi-discrete case (Altschuler, Niles-Weed, Stromme ; 2021) (Delalande ; 2021);
- Convergence rate in the continuous case under strong regularity assumptions (Pal, 2019), second-order expansion for dynamical quadratic OT (Conforti-Tamanini, 2019)
- Convergence rate for $2-$marginal and a general class of $\mathscr{C}^2$ non-degenerate costs (Carlier, Pegon, and Tamanini 2022), more precise estimates by (Malamut, Sylvestre ; 2023)
- Upper bound for the multi-marginal (Eckstein and Nutz 2022) ***with a condition on the optimal transport plans*** in terms of quantization dimension ;

Direction of our work

Goal

- Asymptotics as $\varepsilon \to 0$ (lower and upper bounds) for multi-marginal optimal transport cost;
- Possibly degenerate ground costs c (for 2-marginal, $D^2_{xy}c$ not necessarily invertible);
- No assumption on the optimal transport plans for MOT_0;

Goal

- Asymptotics as $\varepsilon \to 0$ (lower and upper bounds) for multi-marginal optimal transport cost;
- Possibly degenerate ground costs c (for 2-marginal, $D^2_{xy}c$ not necessarily invertible);
- No assumption on the optimal transport plans for MOT_0;

In particular we obtain:

Theorem (Nenna - P.)

Let μ_i be compactly supported measures over X_i with L^∞ densities. Assume that $c \in \mathscr{C}^2(\mathbf{X})$ and satisfies a signature condition on second mixed derivatives. Then

$$\mathrm{MOT}_\varepsilon = \mathrm{MOT}_0 + \frac{1}{2}\left(\sum_{i=1}^{m} d_i - \max_i d_i\right)\varepsilon \log(1/\varepsilon) + O(\varepsilon).$$

Application for ***Wasserstein barycenters*** (more later).

The upper bound

The upper bound

Assumptions

- μ_i are compactly supported measures in $L^\infty(X_i)$ where X_i are $\mathscr{C}^2$ submanifolds of dimension d_i ;
- $c \in \mathscr{C}^{1,1}_{\mathrm{loc}}(\mathsf{X})$ or more generally locally semi-concave (also, weaker upper bound $c \in \mathscr{C}^{0,1}(\mathsf{X})$);

Goal: get an upper bound of the form

$$\mathrm{MOT}_\varepsilon - \mathrm{MOT}_0 \leq \frac{1}{2}\left(\sum_{1\leq i\leq m} d_i - \max_i d_i \right) \varepsilon \log(1/\varepsilon) + O(\varepsilon).$$

Assumptions

- μ_i are compactly supported measures in $L^\infty(X_i)$ where X_i are $\mathscr{C}^2$ submanifolds of dimension d_i ;
- $c \in \mathscr{C}^{1,1}_{\mathrm{loc}}(\mathsf{X})$ or more generally locally semi-concave (also, weaker upper bound $c \in \mathscr{C}^{0,1}(\mathsf{X})$);

Goal: get an upper bound of the form

$$\mathrm{MOT}_\varepsilon - \mathrm{MOT}_0 \leq \frac{1}{2}\Big(\sum_{1\leq i\leq m} d_i - \max_i d_i\Big)\varepsilon \log(1/\varepsilon) + O(\varepsilon).$$

Strategy. Straightforward (almost) generalization of the upper bound in (Carlier, Pegon, and Tamanini 2022) on $\mathscr{C}^2$ submanifolds:

- Build a suitable competitor for the entropic (primal) problem

$$\mathrm{MOT}_\varepsilon = \inf_{\gamma\in\Pi(\mu_1,\dots,\mu_m)} \left\{ \int_{\mathsf{X}} c(x_1,\dots,x_m)\, \mathrm{d}\gamma(x_1,\dots,x_m) + \varepsilon \mathrm{Ent}(\gamma \mid \otimes_{i=1}^m \mu_i) \right\}.$$

using an optimizer for (MOT_0) and a block-approximation of (Carlier, Duval, et al. 2017).

The upper bound

Assumptions

- μ_i are compactly supported measures in $L^\infty(X_i)$ where X_i are $\mathscr{C}^2$ submanifolds of dimension d_i ;
- $c \in \mathscr{C}^{1,1}_{\mathrm{loc}}(\mathrm{X})$ or more generally locally semi-concave (also, weaker upper bound $c \in \mathscr{C}^{0,1}(\mathrm{X})$);

Goal: get an upper bound of the form

$$\mathrm{MOT}_\varepsilon - \mathrm{MOT}_0 \leq \frac{1}{2}\Big(\sum_{1\leq i\leq m} d_i - \max_i d_i\Big)\varepsilon \log(1/\varepsilon) + O(\varepsilon).$$

Strategy. Straightforward (almost) generalization of the upper bound in (Carlier, Pegon, and Tamanini 2022) on $\mathscr{C}^2$ submanifolds:

- Build a suitable competitor for the entropic (primal) problem

$$\mathrm{MOT}_\varepsilon = \inf_{\gamma\in\Pi(\mu_1,\dots,\mu_m)} \left\{ \int_{\mathrm{X}} c(x_1,\dots,x_m)\, \mathrm{d}\gamma(x_1,\dots,x_m) + \varepsilon \mathrm{Ent}(\gamma \,|\, \otimes_{i=1}^m \mu_i) \right\}.$$

  using an optimizer for (MOT_0) and a block-approximation of (Carlier, Duval, et al. 2017).
- Show and use some integral variant of Alexandrov theorem on convex functions.

Upper bound: some details for $m = 2$, marginals μ^-, μ^+

For blocks $\bigsqcup_n A_n = \mathbb{R}^N$ of diameter $\leq \delta$, take as competitor

$$\gamma^\delta := \sum_{i,j\in\mathbb{N}} \gamma_0(A_i \times A_j) \frac{\mu^- \llcorner A_i}{\mu^-(A_i)} \otimes \frac{\mu^+ \llcorner A_j}{\mu^+(A_j)}.$$

Upper bound: some details for $m = 2$, marginals μ^-, μ^+

For blocks $\bigsqcup_n A_n = \mathbb{R}^N$ of diameter $\leq \delta$, take as competitor

$$\gamma^\delta := \sum_{i,j \in \mathbb{N}} \gamma_0(A_i \times A_j) \frac{\mu^- \llcorner A_i}{\mu^-(A_i)} \otimes \frac{\mu^+ \llcorner A_j}{\mu^+(A_j)}.$$

- Plug this competitor into the primal problem, write $E = c - \phi \oplus \psi$ the duality gap, then:

$$\mathrm{MOT}_\varepsilon \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} c \, \mathrm{d}\gamma^\delta + \varepsilon \mathrm{Ent}(\gamma^\delta | \mu^- \otimes \mu^+) = \mathrm{MOT}_0 + \int_{\mathbb{R}^d \times \mathbb{R}^d} E \, \mathrm{d}(\gamma^\delta - \gamma^0) + \varepsilon \mathrm{Ent}(\gamma^\delta | \mu^- \otimes \mu^+)$$

Upper bound: some details for $m = 2$, marginals μ^-, μ^+

For blocks $\bigsqcup_n A_n = \mathbb{R}^N$ of diameter $\leq \delta$, take as competitor

$$\gamma^\delta := \sum_{i,j \in \mathbb{N}} \gamma_0(A_i \times A_j) \frac{\mu^- \llcorner A_i}{\mu^-(A_i)} \otimes \frac{\mu^+ \llcorner A_j}{\mu^+(A_j)}.$$

- Plug this competitor into the primal problem, write $E = c - \phi \oplus \psi$ the duality gap, then:

$$\mathrm{MOT}_\varepsilon \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} c \,\mathrm{d}\gamma^\delta + \varepsilon \mathrm{Ent}(\gamma^\delta | \mu^- \otimes \mu^+) = \mathrm{MOT}_0 + \int_{\mathbb{R}^d \times \mathbb{R}^d} E \,\mathrm{d}(\gamma^\delta - \gamma^0) + \varepsilon \mathrm{Ent}(\gamma^\delta | \mu^- \otimes \mu^+)$$

- Bound the entropy term, for well-chosen blocks:

$$\begin{aligned} \mathrm{Ent}(\gamma^\delta \,|\, \mu^- \otimes \mu^+) &= \sum_{i,j \in \mathbb{N}} \gamma_0(A_i \times A_j) \log\left(\frac{\gamma_0(A_i \times A_j)}{\mu^-(A_i)\mu^+(A_j)} \right) \\ &\leq \sum_{j \in \mathbb{N}} \mu^+(A_j) \log(1/\mu^+(A_j)) = d^+ \log(1/\delta) + O(1). \end{aligned}$$

Upper bound: some details for $m=2$, marginals μ^-, μ^+

For blocks $\bigsqcup_n A_n = \mathbb{R}^N$ of diameter $\leq \delta$, take as competitor

$$\gamma^\delta := \sum_{i,j\in\mathbb{N}} \gamma_0(A_i \times A_j) \frac{\mu^- \llcorner A_i}{\mu^-(A_i)} \otimes \frac{\mu^+ \llcorner A_j}{\mu^+(A_j)}.$$

- Plug this competitor into the primal problem, write $E = c - \phi \oplus \psi$ the duality gap, then:

$$\mathrm{MOT}_\varepsilon \leq \int_{\mathbb{R}^d\times\mathbb{R}^d} c \, \mathrm{d}\gamma^\delta + \varepsilon \mathrm{Ent}(\gamma^\delta|\mu^- \otimes \mu^+) = \mathrm{MOT}_0 + \int_{\mathbb{R}^d\times\mathbb{R}^d} E \, \mathrm{d}(\gamma^\delta - \gamma^0) + \varepsilon \mathrm{Ent}(\gamma^\delta|\mu^- \otimes \mu^+)$$

- Bound the entropy term, for well-chosen blocks:

$$\mathrm{Ent}(\gamma^\delta \,|\, \mu^- \otimes \mu^+) = \sum_{i,j\in\mathbb{N}} \gamma_0(A_i \times A_j) \log\left(\frac{\gamma_0(A_i \times A_j)}{\mu^-(A_i)\mu^+(A_j)}\right)$$

$$\leq \sum_{j\in\mathbb{N}} \mu^+(A_j) \log(1/\mu^+(A_j)) = d^+ \log(1/\delta) + O(1).$$

- Show that $\int_{\mathbb{R}^d\times\mathbb{R}^d} E \, \mathrm{d}(\gamma^\delta - \gamma^0) = O(\delta^2)$ then take $\varepsilon = \delta^2$ (integral Alexandrov-type estimate):

$$\mathrm{MOT}_\varepsilon \leq \mathrm{MOT}_0 + O(\delta^2) + d^+ \varepsilon \log(1/\delta) + O(\varepsilon) = \mathrm{MOT}_0 + \frac{d^+}{2} \varepsilon \log(1/\varepsilon) + O(\varepsilon).$$

The lower bound

A signature condition on the second mixed derivatives

- Consider a $c \in \mathscr{C}^2(\mathsf{X})$ and let P be the set of partitions of $\{1, \ldots, m\}$ into two non empty disjoint subsets: $p := \{p_-, p_+\} \in P$;

A signature condition on the second mixed derivatives

- Consider a $c \in \mathscr{C}^2(\mathsf{X})$ and let P be the set of partitions of $\{1, \ldots, m\}$ into two non empty disjoint subsets: $p := \{p_-, p_+\} \in P$;
- For each $p \in P$ we denote by g_p the bilinear form on $T\mathsf{X}$ as

$$g_p = \sum_{i \in p_-, j \in p_+} D^2_{x_i, x_j} c + D^2_{x_j, x_i} c \quad \text{where} \quad D^2_{x_i, x_j} c = \sum_{\alpha_i, \alpha_j} \frac{\partial^2 c}{\partial_{x_i^{\alpha_i}} \partial_{x_j^{\alpha_j}}} dx_i^{\alpha_i} \otimes dx_j^{\alpha_j}.$$

A signature condition on the second mixed derivatives

- Consider a $c \in \mathscr{C}^2(\mathsf{X})$ and let P be the set of partitions of $\{1, \ldots, m\}$ into two non empty disjoint subsets: $p := \{p_-, p_+\} \in P$;
- For each $p \in P$ we denote by g_p the bilinear form on $T\mathsf{X}$ as

$$g_p = \sum_{i \in p_-, j \in p_+} D^2_{x_i, x_j} c + D^2_{x_j, x_i} c \quad \text{where} \quad D^2_{x_i, x_j} c = \sum_{\alpha_i, \alpha_j} \frac{\partial^2 c}{\partial_{x_i^{\alpha_i}} \partial_{x_j^{\alpha_j}}} dx_i^{\alpha_i} \otimes dx_j^{\alpha_j}.$$

- Define $G := \{\sum_{p \in P} t_p g_p \mid (t_p)_{p \in P} \in \Delta_P\}$ to be the convex hull generated by the g_p.

A signature condition on the second mixed derivatives

- Consider a $c \in \mathscr{C}^2(\mathsf{X})$ and let P be the set of partitions of $\{1, \ldots, m\}$ into two non empty disjoint subsets: $p := \{p_-, p_+\} \in P$;
- For each $p \in P$ we denote by g_p the bilinear form on $T\mathsf{X}$ as

$$g_p = \sum_{i \in p_-, j \in p_+} D^2_{x_i, x_j} c + D^2_{x_j, x_i} c \quad \text{where} \quad D^2_{x_i, x_j} c = \sum_{\alpha_i, \alpha_j} \frac{\partial^2 c}{\partial_{x_i^{\alpha_i}} \partial_{x_j^{\alpha_j}}} dx_i^{\alpha_i} \otimes dx_j^{\alpha_j}.$$

- Define $G := \{\sum_{p \in P} t_p g_p \mid (t_p)_{p \in P} \in \Delta_P\}$ to be the convex hull generated by the g_p.

Theorem (Upper bound on the dimension of the support of the optimal plan (Pass 2011))

Let γ_0 a solution to MOT_0 and suppose that at some point $\mathsf{x} \in \mathsf{X}$, the signature of some $g \in G$ is $(d^+(g), d^-(g), d^0(g))$. Then, there exists a neighbourhood N_{x} of x such that $N_{\mathsf{x}} \bigcap \mathrm{spt}(\gamma_0)$ is contained in a Lipschitz submanifold with dimension no greater than $\sum_i d_i - d^+(g)$.

Lower bound

- μ_i be compactly supported measures over X_i with L^∞ densities;
- $c \in \mathscr{C}^2(\mathsf{X})$;
- for every $\mathsf{x} \in \mathsf{X}$, $\exists g_x \in G, \quad d^+(g_{\mathsf{x}}) \geq d^\star$;

Goal: get a lower bound of the form

$$\mathrm{MOT}_\varepsilon - \mathrm{MOT}_0 \geq \frac{d^\star}{2}\varepsilon \log(1/\varepsilon) - L\varepsilon.$$

Lower bound

- μ_i be compactly supported measures over X_i with L^∞ densities;
- $c \in \mathscr{C}^2(\mathsf{X})$;
- for every $\mathsf{x} \in \mathsf{X}$, $\exists g_x \in G, \quad d^+(g_{\mathsf{x}}) \geq d^\star$;

Goal: get a lower bound of the form

$$\mathrm{MOT}_\varepsilon - \mathrm{MOT}_0 \geq \frac{d^\star}{2}\varepsilon \log(1/\varepsilon) - L\varepsilon.$$

Strategy

- Use the dual regularized problem (in $\log$ form):
- Take Kantorovich potentials (solution to un-regularized dual) as competitors and show that the duality gap $E \doteq c - \oplus_{i=1}^m \phi_i$ grows enough near $\Sigma = \{E = 0\}$.

Lower bound: some details

Let $\{1, \ldots, m\} = p_- \sqcup p_+$, we identify $\boldsymbol{x} \in X$ with (x_-, x_+) and write $\phi_\pm(y) = \sum_{i \in p_\pm} \phi_i(y_i)$.

Lower bound: some details

Let $\{1,\dots,m\} = p_- \sqcup p_+$, we identify $\boldsymbol{x} \in X$ with (x_-, x_+) and write $\phi_\pm(y) = \sum_{i\in p_\pm} \phi_i(y_i)$.

- If (ϕ_i) are c-conjugate, for $\boldsymbol{x}, \boldsymbol{x}' \in X$, we have:

$$\begin{aligned} E(\boldsymbol{x}') &= c(x'_-, x'_+) - \phi_-(x'_-) - \phi_+(x'_+) \\ &\geq c(x'_-, x'_+) - (c(x'_-, x_+) - \phi_+(x_+)) - (c(x_-, x'_+) - \phi_-(x_-)) \\ &= c(x'_-, x'_+) - c(x'_-, x_+) - c(x_-, x'_+) + c(x_-, x_+) - E(\boldsymbol{x}). \end{aligned}$$

Lower bound: some details

Let $\{1,\dots,m\} = p_- \sqcup p_+$, we identify $\boldsymbol{x} \in X$ with (x_-, x_+) and write $\phi_\pm(y) = \sum_{i\in p_\pm} \phi_i(y_i)$.

- If (ϕ_i) are c-conjugate, for $\boldsymbol{x}, \boldsymbol{x}' \in X$, we have:

$$\begin{aligned} E(\boldsymbol{x}') &= c(x'_-, x'_+) - \phi_-(x'_-) - \phi_+(x'_+) \\ &\geq c(x'_-, x'_+) - (c(x'_-, x_+) - \phi_+(x_+)) - (c(x_-, x'_+) - \phi_-(x_-)) \\ &= c(x'_-, x'_+) - c(x'_-, x_+) - c(x_-, x'_+) + c(x_-, x_+) - E(\boldsymbol{x}). \end{aligned}$$

- By Taylor's integral formula

$$E(\boldsymbol{x}') + E(\boldsymbol{x}) \geq \int_0^1 \int_0^1 D^2_{p_- p_+} c(\boldsymbol{x}_{s,t})(x'_- - x_-, x'_+ - x_+) = \frac{1}{2} g_p(\bar{\boldsymbol{x}})(\boldsymbol{x}' - \boldsymbol{x}) + O_{\bar{x}}(\|\boldsymbol{x}' - \boldsymbol{x}\|^2)$$

Let $\{1,\ldots,m\} = p_- \sqcup p_+$, we identify $\boldsymbol{x} \in X$ with (x_-, x_+) and write $\phi_\pm(y) = \sum_{i\in p_\pm} \phi_i(y_i)$.

- If (ϕ_i) are c-conjugate, for $\boldsymbol{x}, \boldsymbol{x'} \in \mathrm{X}$, we have:

$$\begin{aligned} E(\boldsymbol{x'}) &= c(x'_-, x'_+) - \phi_-(x'_-) - \phi_+(x'_+) \\ &\geq c(x'_-, x'_+) - (c(x'_-, x_+) - \phi_+(x_+)) - (c(x_-, x'_+) - \phi_-(x_-)) \\ &= c(x'_-, x'_+) - c(x'_-, x_+) - c(x_-, x'_+) + c(x_-, x_+) - E(\boldsymbol{x}). \end{aligned}$$

- By Taylor's integral formula

$$E(\boldsymbol{x'}) + E(\boldsymbol{x}) \geq \int_0^1 \int_0^1 D^2_{p_- p_+} c(\boldsymbol{x}_{s,t})(x'_- - x_-, x'_+ - x_+) = \frac{1}{2} g_p(\bar{\boldsymbol{x}})(\boldsymbol{x'} - \boldsymbol{x}) + O_{\bar{\boldsymbol{x}}}(\|\boldsymbol{x'} - \boldsymbol{x}\|^2)$$

and taking a convex combination $g = \sum t_p g_p$, for diagonalizing coordinates (u^+, u^-, u^0)

$$E(\boldsymbol{x'}) + E(\boldsymbol{x}) \geq |u^+(\boldsymbol{x'}) - u^+(\boldsymbol{x})|^2 - |u^-(\boldsymbol{x'}) - u^-(\boldsymbol{x})|^2 + O(|\boldsymbol{x'} - \boldsymbol{x}|^2)$$

$\Longrightarrow$ quadratic detachment of the duality gap E in $d^+(g) \geq d^\star$ dimensions.

Lower bound: some details

Let $\{1,\dots,m\} = p_- \sqcup p_+$, we identify $\boldsymbol{x} \in X$ with (x_-, x_+) and write $\phi_\pm(y) = \sum_{i \in p_\pm} \phi_i(y_i)$.

- If (ϕ_i) are c-conjugate, for $\boldsymbol{x}, \boldsymbol{x'} \in X$, we have:

$$\begin{aligned} E(\boldsymbol{x'}) &= c(x'_-, x'_+) - \phi_-(x'_-) - \phi_+(x'_+) \\ &\geq c(x'_-, x'_+) - (c(x'_-, x_+) - \phi_+(x_+)) - (c(x_-, x'_+) - \phi_-(x_-)) \\ &= c(x'_-, x'_+) - c(x'_-, x_+) - c(x_-, x'_+) + c(x_-, x_+) - E(\boldsymbol{x}). \end{aligned}$$

- By Taylor's integral formula

$$E(\boldsymbol{x'}) + E(\boldsymbol{x}) \geq \int_0^1 \int_0^1 D^2_{p_- p_+} c(\boldsymbol{x}_{s,t})(x'_- - x_-, x'_+ - x_+) = \frac{1}{2} g_p(\bar{\boldsymbol{x}})(\boldsymbol{x'} - \boldsymbol{x}) + O_{\bar{x}}(\|\boldsymbol{x'} - \boldsymbol{x}\|^2)$$

  and taking a convex combination $g = \sum t_p g_p$, for diagonalizing coordinates (u^+, u^-, u^0)

$$E(\boldsymbol{x'}) + E(\boldsymbol{x}) \geq |u^+(\boldsymbol{x'}) - u^+(\boldsymbol{x})|^2 - |u^-(\boldsymbol{x'}) - u^-(\boldsymbol{x})|^2 + O(|\boldsymbol{x'} - \boldsymbol{x}|^2)$$

  $\Longrightarrow$ quadratic detachment of the duality gap E in $d^+(g) \geq d^\star$ dimensions.

- Taking $(\phi_i)_{1 \leq i \leq m}$ as competitor in the dual of the entropic MOT:

$$\mathrm{MOT}_\varepsilon \geq \mathrm{MOT}_0 - \varepsilon \log\left(\int_{\Pi_{1 \leq i \leq m} X_i} e^{-\frac{E}{\varepsilon}} \, \mathrm{d} \otimes_{1 \leq i \leq m} \mu_i \right) \geq \mathrm{MOT}_0 + \frac{d^\star}{2} \varepsilon \log(1/\varepsilon) - O(\varepsilon).$$

Consequences and some examples

- $m = 2$ and **non-degenerate** cost function, then we retrieve the bounds in (Carlier, Pegon, and Tamanini 2022);

Consequences and some examples

- $m = 2$ and **non-degenerate** cost function, then we retrieve the bounds in (Carlier, Pegon, and Tamanini 2022);
- $m = 2$ and **degenerate** cost function such that $D^2_{x,y}c$ has rank r then the lower bound is such that $d^\star = r$;

- $m = 2$ and **non-degenerate** cost function, then we retrieve the bounds in (Carlier, Pegon, and Tamanini 2022);
- $m = 2$ and **degenerate** cost function such that $D^2_{x,y}c$ has rank r then the lower bound is such that $d^\star = r$;
- $m = 2$, $d_2 < d_1$ (aka the **unequal dimensional case**) and $D^2_{x,y}c$ has full rank d_2 then

$$\mathrm{MOT}_\varepsilon = \mathrm{MOT}_0 + \frac{d_2}{2}\varepsilon \log(1/\varepsilon) + O(\varepsilon)$$

- $m = 2$ and **non-degenerate** cost function, then we retrieve the bounds in (Carlier, Pegon, and Tamanini 2022);
- $m = 2$ and **degenerate** cost function such that $D^2_{x,y}c$ has rank r then the lower bound is such that $d^\star = r$;
- $m = 2$, $d_2 < d_1$ (aka the **unequal dimensional case**) and $D^2_{x,y}c$ has full rank d_2 then

$$\mathrm{MOT}_\varepsilon = \mathrm{MOT}_0 + \frac{d_2}{2}\varepsilon\log(1/\varepsilon) + O(\varepsilon)$$

- Consider $d_i = d$ for all i and the cost $c = h(\sum_{i=1}^m x_i)$ with $D^2 h < 0$ then $d^* = (m-1)d$ and

$$\mathrm{MOT}_\varepsilon = \mathrm{MOT}_0 + \frac{(m-1)d}{2}\varepsilon\log(1/\varepsilon) + O(\varepsilon).$$

This is the case of Gangbo-Święch cost, that is $\sum_{i<j}|x_i - x_j|^2$ which corresponds to the multi-marginal formulation of the Wasserstein barycenter problem.

Thank You.