An ODE characterisation of entropic (multi-marginal) optimal transport

Luca Nenna

joint work with B. Pass

Numerical methods for optimal transport problems, mean field games and multi-agent dynamics, Universidad Técnica Federico Santa María, 08/01/2024, Valparaíso

(LMO) Université Paris-Saclay and INRIA-Saclay (ParMA)

Overview

1. A crash introduction to (multi-marginal) Optimal Transport
    - Classical Optimal transport
    - Multi-marginal optimal transport
2. Entropic Multi-Marginal Optimal Transport
3. The ODE
4. The algorithm and some numerical results
5. An extension to general (entropic) multi-marginal problem

A crash introduction to (multi-marginal) Optimal Transport

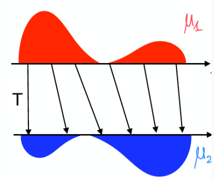
Classical Optimal Transportation Theory

Consider two probability measures μ_i on $X_i \subseteq \mathbb{R}^d$, and c a cost function (e.g. continuous or l.s.c.), the Optimal Transport (OT) problem is defined as follows

$$\mathrm{OT}_0 := \inf \left\{ \int_{\boldsymbol{X}} c(x_1, x_2) \mathrm{d}\gamma(x_1, x_2) \mid \gamma \in \Pi(\mu_1, \mu_2) \right\} \tag{1}$$

where $\Pi(\mu_1, \mu_2)$ denotes the set of couplings $\gamma(x_1, x_2) \in \mathcal{P}(\boldsymbol{X})$ having μ_1 and μ_2 as marginals.

- Solution à la Monge the transport plan γ is deterministic (or à la Monge) if $\gamma = (Id, T)_\sharp \mu$ where $T_\sharp \mu_1 = \mu_2$.



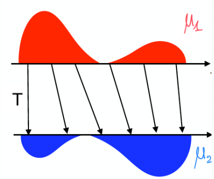
Classical Optimal Transportation Theory

Consider two probability measures μ_i on $X_i \subseteq \mathbb{R}^d$, and c a cost function (e.g. continuous or l.s.c.), the Optimal Transport (OT) problem is defined as follows

$$\mathrm{OT}_0 := \inf \left\{ \int_{\boldsymbol{X}} c(x_1, x_2) \mathrm{d}\gamma(x_1, x_2) \mid \gamma \in \Pi(\mu_1, \mu_2) \right\} \tag{1}$$

where $\Pi(\mu_1, \mu_2)$ denotes the set of couplings $\gamma(x_1, x_2) \in \mathcal{P}(\boldsymbol{X})$ having μ_1 and μ_2 as marginals.

- Solution à la Monge the transport plan γ is deterministic (or à la Monge) if $\gamma = (Id, T)_\sharp \mu$ where $T_\sharp \mu_1 = \mu_2$.



- Duality:

$$\sup \{ \mathcal{J}(\phi_1, \phi_2) \mid (\phi_1, \phi_2) \in \mathcal{K} \} . \tag{2}$$

where

$$\mathcal{J}(\phi_1, \phi_2) := \int_{X_1} \phi_1 \mathrm{d}\mu_1 + \int_{X_2} \phi_2 \mathrm{d}\mu_2$$

and $\mathcal{K}$ is the set of bounded and continuous functions (ϕ_1, ϕ_2) such that $\phi_1(x_1) + \phi(x_2) \leq c(x_1, x_2)$.

The Multi-Marginal Optimal Transportation

Take **(1)** m probability measures $\mu_i \in \mathcal{P}(X_i)$; **(2)** c a cost function. Then the multi-marginal OT problem reads as:

The Multi-Marginal Optimal Transportation

Take **(1)** m probability measures $\mu_i \in \mathcal{P}(X_i)$; **(2)** c a cost function. Then the multi-marginal OT problem reads as:

Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_0 := \inf_{\gamma \in \Pi(\mu_1, \ldots, \mu_m)} \int_{\boldsymbol{X}} c(x_1, \ldots, x_m) \mathrm{d}\gamma(x_1, \ldots, x_m) \tag{3}$$

where $\Pi(\mu_1, \ldots, \mu_m)$ denotes the set of couplings $\gamma(x_1, \ldots, x_m)$ having μ_i as marginals.

The Multi-Marginal Optimal Transportation

Take **(1)** m probability measures $\mu_i \in \mathcal{P}(X_i)$; **(2)** c a cost function. Then the multi-marginal OT problem reads as:

Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_0 := \inf_{\gamma \in \Pi(\mu_1,\dots,\mu_m)} \int_{\boldsymbol{X}} c(x_1,\dots,x_m)\mathrm{d}\gamma(x_1,\dots,x_m) \tag{3}$$

where $\Pi(\mu_1,\dots,\mu_m)$ denotes the set of couplings $\gamma(x_1,\dots,x_m)$ having μ_i as marginals.

- Solution à la Monge: $\gamma = (Id, T_2, \dots, T_m)_\sharp \mu_1$ where $T_{i\sharp}\mu_1 = \mu_i$.

The Multi-Marginal Optimal Transportation

Take **(1)** m probability measures $\mu_i \in \mathcal{P}(X_i)$; **(2)** c a cost function. Then the multi-marginal OT problem reads as:

Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_0 := \inf_{\gamma \in \Pi(\mu_1, \ldots, \mu_m)} \int_{\boldsymbol{X}} c(x_1, \ldots, x_m) \mathrm{d}\gamma(x_1, \ldots, x_m) \tag{3}$$

where $\Pi(\mu_1, \ldots, \mu_m)$ denotes the set of couplings $\gamma(x_1, \ldots, x_m)$ having μ_i as marginals.

- Solution à la Monge: $\gamma = (Id, T_2, \ldots, T_m)_\sharp \mu_1$ where $T_{i\sharp}\mu_1 = \mu_i$.
- Duality: Both 2 and m marginal OT problems admit a useful dual formulation

Take **(1)** m probability measures $\mu_i \in \mathcal{P}(X_i)$; **(2)** c a cost function. Then the multi-marginal OT problem reads as:

Multi-Marginal Optimal Transport problem

It reads as:

$$\mathrm{MOT}_0 := \inf_{\gamma \in \Pi(\mu_1, \dots, \mu_m)} \int_{\boldsymbol{X}} c(x_1, \dots, x_m) \mathrm{d}\gamma(x_1, \dots, x_m) \quad (3)$$

where $\Pi(\mu_1, \dots, \mu_m)$ denotes the set of couplings $\gamma(x_1, \dots, x_m)$ having μ_i as marginals.

- Solution à la Monge: $\gamma = (Id, T_2, \dots, T_m)_\sharp \mu_1$ where $T_{i\sharp}\mu_1 = \mu_i$.
- Duality: Both 2 and m marginal OT problems admit a useful dual formulation

Why is it a difficult problem to treat?
Example: $m = 3$, $d = 1$, $\mu_i = \mathcal{L}_{[0,1]}$ $\forall i$ and $c(x_1, x_2, x_3) = |x_1 + x_2 + x_3|^2$.

- Uniqueness fails **(Simone Di Marino, Gerolin, and Luca Nenna 2017)**;
- $\exists$ T_i optimal, are not differentiable at any point and they are fractal maps **ibid., Thm 4.6**

Why are we interested in MOT?

- The Wasserstein barycenter problem can be rewritten as a MMOT problem (see **(Agueh and G. Carlier 2011)**): statistics, machine learning, image processing;

Why are we interested in MOT?

- The Wasserstein barycenter problem can be rewritten as a MMOT problem (see **(Agueh and G. Carlier 2011)**): statistics, machine learning, image processing;
- Matching for teams problem (see **(Guillaume Carlier and Ekeland 2010)**): economics.

Why are we interested in MOT?

- The Wasserstein barycenter problem can be rewritten as a MMOT problem (see **(Agueh and G. Carlier 2011)**): statistics, machine learning, image processing;
- Matching for teams problem (see **(Guillaume Carlier and Ekeland 2010)**): economics.
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Gori-Giorgi 2012; Cotar, Friesecke, and Klüppelberg 2013)**). The plan $\gamma(x_1, \ldots, x_m)$ returns the probability of finding electrons at position $x_1, \ldots, x_m$;

Why are we interested in MOT?

- The Wasserstein barycenter problem can be rewritten as a MMOT problem (see **(Agueh and G. Carlier 2011)**): statistics, machine learning, image processing;
- Matching for teams problem (see **(Guillaume Carlier and Ekeland 2010)**): economics.
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Gori-Giorgi 2012; Cotar, Friesecke, and Klüppelberg 2013)**). The plan $\gamma(x_1, \ldots, x_m)$ returns the probability of finding electrons at position $x_1, \ldots, x_m$;
- Incompressible Euler Equations **(Brenier 1989)** : $\gamma(\omega)$ gives "the mass of fluid" which follows a path ω. See also **(Jean-David Benamou, Guillaume Carlier, and Luca Nenna 2018)**.

Why are we interested in MOT?

- The Wasserstein barycenter problem can be rewritten as a MMOT problem (see **(Agueh and G. Carlier 2011)**): statistics, machine learning, image processing;
- Matching for teams problem (see **(Guillaume Carlier and Ekeland 2010)**): economics.
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Gori-Giorgi 2012; Cotar, Friesecke, and Klüppelberg 2013)**). The plan $\gamma(x_1, \dots, x_m)$ returns the probability of finding electrons at position $x_1, \dots, x_m$;
- Incompressible Euler Equations **(Brenier 1989)** : $\gamma(\omega)$ gives "the mass of fluid" which follows a path ω. See also **(Jean-David Benamou, Guillaume Carlier, and Luca Nenna 2018)**.
- Mean Field Games **(J.-D. Benamou, G. Carlier, S. Di Marino, and L. Nenna 2018)**;

Why are we interested in MOT?

- The Wasserstein barycenter problem can be rewritten as a MMOT problem (see **(Agueh and G. Carlier 2011)**): statistics, machine learning, image processing;
- Matching for teams problem (see **(Guillaume Carlier and Ekeland 2010)**): economics.
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Gori-Giorgi 2012; Cotar, Friesecke, and Klüppelberg 2013)**). The plan $\gamma(x_1, \dots, x_m)$ returns the probability of finding electrons at position $x_1, \dots, x_m$;
- Incompressible Euler Equations **(Brenier 1989)** : $\gamma(\omega)$ gives "the mass of fluid" which follows a path ω. See also **(Jean-David Benamou, Guillaume Carlier, and Luca Nenna 2018)**.
- Mean Field Games **(J.-D. Benamou, G. Carlier, S. Di Marino, and L. Nenna 2018)**;
- Risk measures **(Ennaji, Mérigot, Luca Nenna, and Pass 2022)**

Why are we interested in MOT?

- The Wasserstein barycenter problem can be rewritten as a MMOT problem (see **(Agueh and G. Carlier 2011)**): statistics, machine learning, image processing;
- Matching for teams problem (see **(Guillaume Carlier and Ekeland 2010)**): economics.
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Gori-Giorgi 2012; Cotar, Friesecke, and Klüppelberg 2013)**). The plan $\gamma(x_1, \dots, x_m)$ returns the probability of finding electrons at position $x_1, \dots, x_m$;
- Incompressible Euler Equations **(Brenier 1989)** : $\gamma(\omega)$ gives "the mass of fluid" which follows a path ω. See also **(Jean-David Benamou, Guillaume Carlier, and Luca Nenna 2018)**.
- Mean Field Games **(J.-D. Benamou, G. Carlier, S. Di Marino, and L. Nenna 2018)**;
- Risk measures **(Ennaji, Mérigot, Luca Nenna, and Pass 2022)**
- Martingale transport (JD's talk), etc

Entropic Multi-Marginal Optimal Transport

Definition of the problem

Consider **(1)** m probability measures μ_i on $X_i \subseteq \mathbb{R}^d$ of dimension d_i; **(2)** a cost function $c : \boldsymbol{X} \to \mathbb{R}_+$ (e.g. continuous or lsc) where $\boldsymbol{X} := \times_i^m X_i$;

Consider **(1)** m probability measures μ_i on $X_i \subseteq \mathbb{R}^d$ of dimension d_i; **(2)** a cost function $c : \boldsymbol{X} \to \mathbb{R}_+$ (e.g. continuous or lsc) where $\boldsymbol{X} := \times_i^m X_i$;

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 the infimum is taken among all *couplings* γ having μ_i as marginals ($\gamma \in \Pi(\mu_1,\dots,\mu_m)$), and $\varepsilon > 0$ is a small temperature parameter.

Some remarks:

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

Some remarks:

- $m = 2$. Classical Entropic Optimal Transport.
- $\varepsilon = 0$ and $m = 2$. Classical Monge-Kantorovich problem with value MOT_0: convex problem, but may have several solutions γ, $\mathrm{Ent}(\gamma \mid \otimes_{i=1}^{m} \mu_i)$ may be finite or not!

Some remarks:

- $m = 2$. Classical Entropic Optimal Transport.
- $\varepsilon = 0$ and $m = 2$. Classical Monge-Kantorovich problem with value MOT_0: convex problem, but may have several solutions γ, $\mathrm{Ent}(\gamma \mid \otimes_{i=1}^m \mu_i)$ may be finite or not!
- $\varepsilon > 0$. Strictly convex cost $\implies$ unique solution $\gamma_\varepsilon = e^{-c/\varepsilon} \prod_{i=1}^m e^{\phi_i/\varepsilon}$ with finite entropy where the ϕ_i are the optimal dual variables.

Some remarks:

- $m = 2$. Classical Entropic Optimal Transport.
- $\varepsilon = 0$ and $m = 2$. Classical Monge-Kantorovich problem with value MOT_0: convex problem, but may have several solutions γ, $\mathrm{Ent}(\gamma \mid \otimes_{i=1}^m \mu_i)$ may be finite or not!
- $\varepsilon > 0$. Strictly convex cost $\implies$ unique solution $\boldsymbol{\gamma_\varepsilon = e^{-c/\varepsilon} \prod_{i=1}^m e^{\phi_i/\varepsilon}}$ with finite entropy where the ϕ_i are the optimal dual variables.
- Asymptotics as $\varepsilon \to 0$

Theorem ((Luca Nenna and Pegon 2023))

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

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

What are we interested in and direction of our work

We are interested in solving the entropic multi-marginal optimal transport.

What are we interested in and direction of our work

We are interested in solving the entropic multi-marginal optimal transport.

Main steps of the work:

1. Introduce a suitable one parameter family of cost functions c_η, **interpolating** between the original multi-marginal problem and a simpler one whose complexity **scales linearly** in the number of marginals;

What are we interested in and direction of our work

We are interested in solving the entropic multi-marginal optimal transport.

Main steps of the work:

1. Introduce a suitable one parameter family of cost functions c_η, **interpolating** between the original multi-marginal problem and a simpler one whose complexity **scales linearly** in the number of marginals;
2. Differentiate the **optimality condition of the dual** $\mathrm{MOT}_\varepsilon := \sup_\phi \tilde{\Phi}(\phi, \eta)$ with respect to η (ε is now fixed);

What are we interested in and direction of our work

We are interested in solving the entropic multi-marginal optimal transport.

Main steps of the work:

1. Introduce a suitable one parameter family of cost functions c_η, **interpolating** between the original multi-marginal problem and a simpler one whose complexity **scales linearly** in the number of marginals;
2. Differentiate the **optimality condition of the dual** $\mathrm{MOT}_\varepsilon := \sup_\phi \tilde{\Phi}(\phi, \eta)$ with respect to η (ε is now fixed);
3. The solution of the original multi-marginal problem can be now recovered by solving an **ordinary differential equation** (ODE) whose initial condition is the solution to the simpler problem;

$$\begin{cases} \dfrac{\mathrm{d}\phi}{\mathrm{d}\eta}(\eta) = -[D^2_{\phi,\phi}\tilde{\Phi}(\phi(\eta), \eta)]^{-1} \dfrac{\partial}{\partial \eta} \nabla_\phi \tilde{\Phi}(\phi(\eta), \eta), \\ \phi(0) = \phi_w, \end{cases}$$

What are we interested in and direction of our work

We are interested in solving the entropic multi-marginal optimal transport.

Main steps of the work:

1. Introduce a suitable one parameter family of cost functions c_η, **interpolating** between the original multi-marginal problem and a simpler one whose complexity **scales linearly** in the number of marginals;
2. Differentiate the **optimality condition of the dual** $\mathrm{MOT}_\varepsilon := \sup_\phi \tilde{\Phi}(\phi, \eta)$ with respect to η (ε is now fixed);
3. The solution of the original multi-marginal problem can be now recovered by solving an **ordinary differential equation** (ODE) whose initial condition is the solution to the simpler problem;

$$\begin{cases} \dfrac{\mathrm{d}\phi}{\mathrm{d}\eta}(\eta) = -[D^2_{\phi,\phi}\tilde{\Phi}(\phi(\eta),\eta)]^{-1}\dfrac{\partial}{\partial\eta}\nabla_\phi\tilde{\Phi}(\phi(\eta),\eta), \\ \phi(0) = \phi_w, \end{cases}$$

Remark: This method is actually inspired by the one introduced in **(G. Carlier, Galichon, and Santambrogio 2009/10)** to compute the Monge solution of the two marginal problem, starting from the Knothe-Rosenblatt rearrangement.

The ODE

How to derive the differential equation

Some assumptions to make it simple:

1. **(Equal marginals and discrete set)** All the marginals are equal $\mu_i = \rho = \sum_{x \in X} \rho_x \delta_x$, where X is a finite subset.
2. **(Pair-wise cost)** $c_\eta(x_1, \ldots, x_m) := \eta \sum_{i=2}^{m} \sum_{j=i+1}^{m} w(x_i, x_j) + \boldsymbol{\sum_{i=2}^{m} w(x_1, x_i)}$.
3. **(Symmetric cost)** The two body cost w is symmetric $w(x, y) = w(x, y)$.
4. **(Finite cost)** The two body cost function $w : X \times X \to \mathbb{R}$ is everywhere real-valued.

How to derive the differential equation

Some assumptions to make it simple:

1. **(Equal marginals and discrete set)** All the marginals are equal $\mu_i = \rho = \sum_{x \in X} \rho_x \delta_x$, where X is a finite subset.
2. **(Pair-wise cost)** $c_\eta(x_1, \dots, x_m) := \eta \sum_{i=2}^{m} \sum_{j=i+1}^{m} w(x_i, x_j) + \boldsymbol{\sum_{i=2}^{m} w(x_1, x_i)}$.
3. **(Symmetric cost)** The two body cost w is symmetric $w(x, y) = w(x, y)$.
4. **(Finite cost)** The two body cost function $w : X \times X \to \mathbb{R}$ is everywhere real-valued.

Step 1: Consider the dual problem (it is convex!);

$$\inf_{\phi} \left\{ \tilde{\Phi}(\phi, \eta) \right\}, \tag{4}$$

where

$$\tilde{\Phi}(\phi, \eta) := -(m-1) \int_X \phi \mathrm{d}\rho + \varepsilon \int_X \underbrace{\log \left(\int_{X^{m-1}} \exp \left(\frac{\sum_{i=2}^{m} \phi - c_\eta}{\varepsilon} \right) \mathrm{d} \otimes^{m-1} \rho \right)}_{\text{Log-Sum-Exp}} \mathrm{d}\rho.$$

Step 2: Thanks to convexity we have that the minimizers are characterized by $\nabla_\phi \tilde{\Phi}(\phi, \eta) = 0$. Then, by differentiate w.r.t. η we obtain

$$\frac{\mathrm{d}\phi}{\mathrm{d}\eta}(\eta) = -[D^2_{\phi,\phi}\tilde{\Phi}(\phi(\eta), \eta)]^{-1} \frac{\partial}{\partial \eta} \nabla_\phi \tilde{\Phi}(\phi(\eta), \eta).$$

Step 2: Thanks to convexity we have that the minimizers are characterized by $\nabla_\phi \tilde{\Phi}(\phi, \eta) = 0$. Then, by differentiate w.r.t. η we obtain

$$\frac{\mathrm{d}\phi}{\mathrm{d}\eta}(\eta) = -[D^2_{\phi,\phi}\tilde{\Phi}(\phi(\eta), \eta)]^{-1} \frac{\partial}{\partial \eta} \nabla_\phi \tilde{\Phi}(\phi(\eta), \eta).$$

Step 3: The following well-posedness theorem then holds.

Theorem

Let $\phi(\eta)$ be the solution to the dual problem above for all $\eta \in [0, 1]$. Then $\eta \mapsto \phi(\eta)$ is $\mathcal{C}^1$ and is the unique solution to the Cauchy problem with $\phi(0) = \phi_w$.

Sketch of the proof:

Step 2: Thanks to convexity we have that the minimizers are characterized by $\nabla_\phi \tilde{\Phi}(\phi, \eta) = 0$. Then, by differentiate w.r.t. η we obtain

$$\frac{\mathrm{d}\phi}{\mathrm{d}\eta}(\eta) = -[D^2_{\phi,\phi}\tilde{\Phi}(\phi(\eta), \eta)]^{-1} \frac{\partial}{\partial \eta} \nabla_\phi \tilde{\Phi}(\phi(\eta), \eta).$$

Step 3: The following well-posedness theorem then holds.

Theorem

Let $\phi(\eta)$ be the solution to the dual problem above for all $\eta \in [0, 1]$. Then $\eta \mapsto \phi(\eta)$ is $\mathcal{C}^1$ and is the unique solution to the Cauchy problem with $\phi(0) = \phi_w$.

Sketch of the proof:

- The pure second derivatives with respect to ϕ as well as the mixed second derivatives with respect to ϕ and η exist and are Lipschitz;

Step 2: Thanks to convexity we have that the minimizers are characterized by $\nabla_\phi \tilde{\Phi}(\phi, \eta) = 0$. Then, by differentiate w.r.t. η we obtain

$$\frac{\mathrm{d}\phi}{\mathrm{d}\eta}(\eta) = -[D^2_{\phi,\phi}\tilde{\Phi}(\phi(\eta), \eta)]^{-1} \frac{\partial}{\partial \eta} \nabla_\phi \tilde{\Phi}(\phi(\eta), \eta).$$

Step 3: The following well-posedness theorem then holds.

Theorem

Let $\phi(\eta)$ be the solution to the dual problem above for all $\eta \in [0,1]$. Then $\eta \mapsto \phi(\eta)$ is $\mathcal{C}^1$ and is the unique solution to the Cauchy problem with $\phi(0) = \phi_w$.

Sketch of the proof:

- The pure second derivatives with respect to ϕ as well as the mixed second derivatives with respect to ϕ and η exist and are Lipschitz;
- The Hessian with respect to ϕ is invertible: since the cost is bounded then the potentials are bounded too (**(Guillaume Carlier 2021)**). So one can restrict the study of the well-posedness of the ODE on the set

$$U := \{\phi \mid \phi_{x_0} = 0, \ ||\phi||_\infty \leq C\}.$$

On this set the functional $\tilde{\Phi}$ is now **strongly convex**.

The algorithm and some numerical results

The algorithm to compute the ODE solution

- Algorithm to compute the ϕ via explicit Euler method takes the following form:

Require: $\phi(0) = \phi_w$

while $||\phi^{(k+1)} - \phi^{(k)}|| < \text{tol}$ **do**

  $D^{(k)} := D^2_{\phi,\phi}\tilde{\Phi}(\phi^{(k)}, kh)$

  $b^{(k)} := -\frac{\partial}{\partial \epsilon}\nabla_{\phi}\tilde{\Phi}(\phi^{(k)}, kh)$

  Solve $D^{(k)}z = b^{(k)}$

  $\phi^{(k+1)} = \phi^{(k)} + hz$

end while

The algorithm to compute the ODE solution

- Algorithm to compute the ϕ via explicit Euler method takes the following form:

Require: $\phi(0) = \phi_w$

while $||\phi^{(k+1)} - \phi^{(k)}|| <$tol **do**

$\quad D^{(k)} := D^2_{\phi,\phi}\tilde{\Phi}(\phi^{(k)}, kh)$

$\quad b^{(k)} := -\frac{\partial}{\partial \epsilon}\nabla_{\phi}\tilde{\Phi}(\phi^{(k)}, kh)$

$\quad$ Solve $D^{(k)}z = b^{(k)}$

$\quad \phi^{(k+1)} = \phi^{(k)} + hz$

end while

Remarks:

- The Euler scheme converges linearly and the uniform error between the discretized solution obtained via the scheme and the solution to the ODE is $O(h)$;
- Thanks to the regularity of the RHS of the ODE one can apply high order methods.

The algorithm to compute the ODE solution

- Algorithm to compute the ϕ via explicit Euler method takes the following form:

Require: $\phi(0) = \phi_w$

1: **while** $||\phi^{(k+1)} - \phi^{(k)}|| <$tol **do**
2: $\quad D^{(k)} := D^2_{\phi,\phi}\tilde{\Phi}(\phi^{(k)}, kh)$
3: $\quad b^{(k)} := -\dfrac{\partial}{\partial \epsilon}\nabla_\phi \tilde{\Phi}(\phi^{(k)}, kh)$
4: $\quad$ Solve $D^{(k)}z = b^{(k)}$
5: $\quad \phi^{(k+1)} = \phi^{(k)} + hz$
6: **end while**

Remarks:

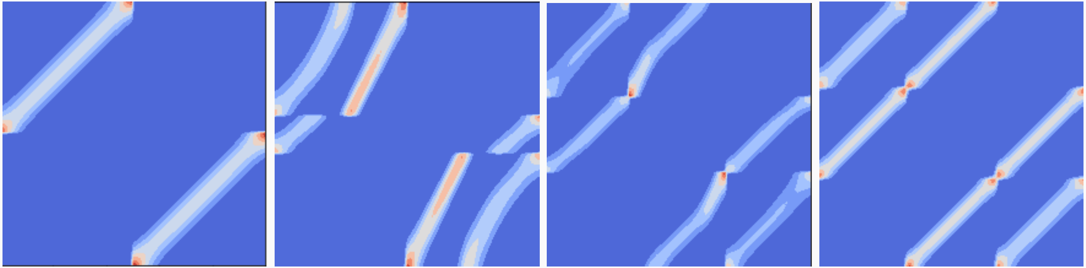
- The Euler scheme converges linearly and the uniform error between the discretized solution obtained via the scheme and the solution to the ODE is $O(h)$;
- Thanks to the regularity of the RHS of the ODE one can apply high order methods.
- At each step k we obtaine the solution of the entropic multi-marginal problem with cost c_{kh}!

Comparison with Sinkhorn

Consider $\varepsilon = 0.006$, $m = 3$, the uniform measure on $[0, 1]$ uniformily discretized with 400 gridpoints, the pairwise interaction $w(x, y) = -\log(0.1 + |x - y|)$ and a reference solution ϕ_ε computed via a gradient descent algorithm. Then we have the following comparison between the ODE approach and Sinkhorn in terms of performances

	3rd RK	5th RK	8th RK	Sinkhorn
relative error	1.47×10^{-5}	7.8×10^{-6}	7.62×10^{-6}	5.46×10^{-6}
iterations	87	87	87	820
CPU time (sec)	72.39	158.9	385.1	102.8

- Log cost and support of the coupling $\gamma^{\eta}_{1,2}$.



$\eta = 0$ $\eta = 0.25$ $\eta = 0.75$ $\eta = 1$

Generalized solutions to incompressible Euler Equations

- Brenier's relaxed formulation consists in finding a probability measure over absolutely continuous paths which minimizes the average kinetic energy.

Generalized solutions to incompressible Euler Equations

- Brenier's relaxed formulation consists in finding a probability measure over absolutely continuous paths which minimizes the average kinetic energy.
- The **incompressibility** at each time t, the distribution of position need be uniform.

- Brenier's relaxed formulation consists in finding a probability measure over absolutely continuous paths which minimizes the average kinetic energy.
- The **incompressibility** at each time t, the distribution of position need be uniform.
- If we consider a uniform discretization of $[0, T]$ (where T is the final time) with m steps in time, we recover a multi-marginal formulation of the Brenier principle with the specific cost function

$$c(x_1, \dots, x_m) = \frac{m^2}{2T^2} \sum_{i=1}^{m-1} |x_{i+1} - x_i|^2 + \beta |F(x_1) - x_m|^2,$$

where $\beta > 0$ is a penalization parameter in order to enforce the initial-final constraint.

- Brenier's relaxed formulation consists in finding a probability measure over absolutely continuous paths which minimizes the average kinetic energy.
- The **incompressibility** at each time t, the distribution of position need be uniform.
- If we consider a uniform discretization of $[0, T]$ (where T is the final time) with m steps in time, we recover a multi-marginal formulation of the Brenier principle with the specific cost function

$$c(x_1, \dots, x_m) = \frac{m^2}{2T^2} \sum_{i=1}^{m-1} |x_{i+1} - x_i|^2 + \beta |F(x_1) - x_m|^2,$$

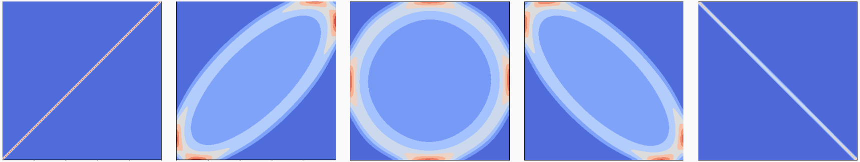
where $\beta > 0$ is a penalization parameter in order to enforce the initial-final constraint.
- If we consider now the ODE setting, we have now to deal with a non symmetric case and so to solve a system, still well posed, of ODEs. In particular we consider the following c_η cost

$$\boxed{c_\eta(x_1, \dots, x_m) = \frac{m^2}{2T^2} |x_2 - x_1|^2 + \eta \left(\frac{m^2}{2T^2} \sum_{i=2}^{m-1} |x_{i+1} - x_i|^2 \right) + \beta |F(x_1) - x_m|^2.}$$

At $\eta = 1$ we plot the coupling $\gamma_{1,i}$ giving the probability of finding a generalized particle initially at x_1 to be at x_i at time i.

At $\eta = 1$ we plot the coupling $\gamma_{1,i}$ giving the probability of finding a generalized particle initially at x_1 to be at x_i at time i.

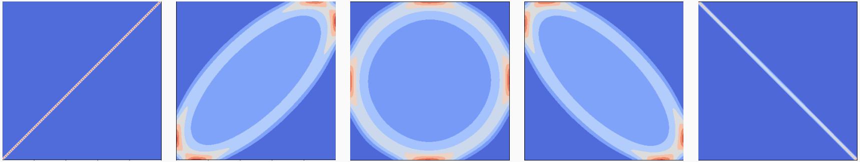
- $F(x) = 1 - x$



$t = 0$ $t = 1/4$ $t = 1/2$ $t = 3/4$ $t = 1$

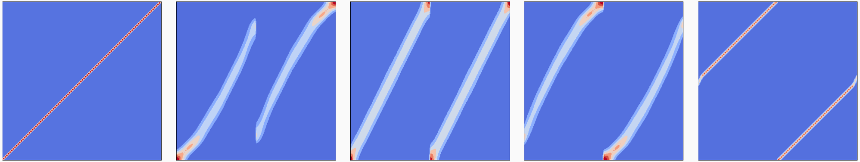
At $\eta = 1$ we plot the coupling $\gamma_{1,i}$ giving the probability of finding a generalized particle initially at x_1 to be at x_i at time i.

- $F(x) = 1 - x$



$t = 0$ $t = 1/4$ $t = 1/2$ $t = 3/4$ $t = 1$

- $F(x) = (x + 1/2) \bmod 1$



$t = 0$ $t = 1/4$ $t = 1/2$ $t = 3/4$ $t = 1$

An extension to general (entropic) multi-marginal problem

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "1st" generalization

$$\boxed{\mathrm{MOT}_\varepsilon := \inf_{\gamma\in\Pi(\mu_1,\ldots,\mu_m)} \left\{ \int_X c(\boldsymbol{\eta}, x_1,\ldots,x_m)\,\mathrm{d}\gamma(x_1,\ldots,x_m) + \varepsilon \mathrm{Ent}(\gamma \mid \otimes_{i=1}^m \mu_i) \right\},}$$

where the cost function is not anymore symmetric but such that $c(0, x_1, \ldots, x_m)$ give a MOT easy to solve:

1. $c(0, x_1, \ldots, x_m) = 0$;

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "1st" generalization

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

where the cost function is not anymore symmetric but such that $c(0, x_1, \ldots, x_m)$ give a MOT easy to solve:

1. $c(0, x_1, \ldots, x_m) = 0$;
2. $c(\eta, x_1, \ldots, x_m)$ is the Euler cost;

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "1st" generalization

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

where the cost function is not anymore symmetric but such that $c(0, x_1, \dots, x_m)$ give a MOT easy to solve:

1. $c(0, x_1, \dots, x_m) = 0$;
2. $c(\eta, x_1, \dots, x_m)$ is the Euler cost;
3. $c(\eta, x_1, z, x_2) = (1-\eta)|x_1 - z|^2 + \eta|z - x_3|^2$, γ is a 3 marginals coupling with only two fixed marginals, μ_1 and μ_2. Then the $z-$marginal of γ gives the **Wasserstein geodesic** at time η.

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "1st" generalization

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

where the cost function is not anymore symmetric but such that $c(0, x_1, \ldots, x_m)$ give a MOT easy to solve:

1. $c(0, x_1, \ldots, x_m) = 0$;
2. $c(\eta, x_1, \ldots, x_m)$ is the Euler cost;
3. $c(\eta, x_1, z, x_2) = (1-\eta)|x_1 - z|^2 + \eta|z - x_3|^2$, γ is a 3 marginals coupling with only two fixed marginals, μ_1 and μ_2. Then the $z-$marginal of γ gives the **Wasserstein geodesic** at time η.
4. $c(\eta, x_1, \ldots, x_m, z) = \sum_{i=1}^m \lambda_i(\eta)|x_i - z|^2$ such that $\sum_{i=1}^m \lambda_i(\eta) = 1$ for every η and γ is an $m+1$ coupling with m fixed marginals. Then at for every η the $z-$marginal of γ is the **Wasserstein barycenter** with weights $\lambda_i(\eta)$.

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "2nd" generalization

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

where $\boldsymbol{\Pi^Q(\mu_1, \ldots, \mu_m)}$ is the set of coupling having $\mu_1, \ldots, \mu_m$ as marginals and satisfying an additional constraint $\int q \mathrm{d}\gamma = 0$ for all $q \in Q$ where Q be a set of bounded continuous function on $\boldsymbol{X}$.

- **Classical case:** $\boldsymbol{Q} = \{0\}$;

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "2nd" generalization

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

where $\boldsymbol{\Pi^Q(\mu_1,\dots,\mu_m)}$ is the set of coupling having $\mu_1, \dots, \mu_m$ as marginals and satisfying an additional constraint $\int q \mathrm{d}\gamma = 0$ for all $q \in Q$ where Q be a set of bounded continuous function on $\boldsymbol{X}$.

- **Classical case:** $\boldsymbol{Q} = \{0\}$;
- **Generalized Euler solution:** force $\gamma_{1,m} = (Id, F)_\sharp \mathcal{L}$;

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "2nd" generalization

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

where $\boldsymbol{\Pi}^{\boldsymbol{Q}}(\boldsymbol{\mu_1,\dots,\mu_m})$ is the set of coupling having $\mu_1,\dots,\mu_m$ as marginals and satisfying an additional constraint $\int q\mathrm{d}\gamma = 0$ for all $q \in Q$ where Q be a set of bounded continuous function on $\boldsymbol{X}$.

- **Classical case:** $\boldsymbol{Q} = \{0\}$;
- **Generalized Euler solution:** force $\gamma_{1,m} = (Id, F)_\sharp \mathcal{L}$;
- **Martingale OT:** $\Pi^Q(\mu_1,\mu_2)$ with extra constraint

$$\int q(x_1)(x_2 - x_1)\mathrm{d}\gamma = 0, \quad \forall \boldsymbol{q} \in \mathcal{C}_b(X_1).$$

Extension to general multi-marginal problems (joint work with B. Pass and J. Zoen-Git Hiew)

Consider the following "2nd" generalization

$$\boxed{\mathrm{MOT}_\varepsilon := \inf_{\gamma \in \boldsymbol{\Pi}^{\boldsymbol{Q}}(\boldsymbol{\mu}_1,\dots,\boldsymbol{\mu}_m)} \left\{ \int_X c(\eta, x_1, \dots, x_m)\, \mathrm{d}\gamma(x_1, \dots, x_m) + \varepsilon \mathrm{Ent}(\gamma \,|\, \otimes_{i=1}^m \mu_i) \right\},}$$

where $\boldsymbol{\Pi}^{\boldsymbol{Q}}(\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_m)$ is the set of coupling having $\mu_1, \dots, \mu_m$ as marginals and satisfying an additional constraint $\int q \mathrm{d}\gamma = 0$ for all $q \in Q$ where Q be a set of bounded continuous function on $\boldsymbol{X}$.

- **Classical case:** $\boldsymbol{Q} = \{0\}$;
- **Generalized Euler solution:** force $\gamma_{1,m} = (Id, F)_\sharp \mathcal{L}$;
- **Martingale OT:** $\Pi^Q(\mu_1, \mu_2)$ with extra constraint

$$\int q(x_1)(x_2 - x_1)\mathrm{d}\gamma = 0, \quad \forall q \in \mathcal{C}_b(X_1).$$

- **Multi-period martingale OT:** e.g. 3−period $\Pi^Q(\mu_1, \mu_2, \mu_3)$ with extra constraint

$$\int [q(x_1)(x_2 - x_1) + h(x_1, x_2)(x_3 - x_2)]\mathrm{d}\gamma = 0, \quad \forall q \in \mathcal{C}_b(X_1), \forall h \in \mathcal{C}_b(X_1 \times X_2).$$

Conclusion

Take-home messages:

- An ODE to characterize entropic multi-marginal optimal transport;

Conclusion

Take-home messages:

- An ODE to characterize entropic multi-marginal optimal transport;
- It works for symmetric and non symmetric cost;

Conclusion

Take-home messages:

- An ODE to characterize entropic multi-marginal optimal transport;
- It works for symmetric and non symmetric cost;
- Regularity allows to use high order methods;

Conclusion

Take-home messages:

- An ODE to characterize entropic multi-marginal optimal transport;
- It works for symmetric and non symmetric cost;
- Regularity allows to use high order methods;
- It allows to interpolate between different costs and for each $\eta \in [0, 1]$ it returns the solution to the corresponding multi-marginal problems.

Conclusion

Take-home messages:

- An ODE to characterize entropic multi-marginal optimal transport;
- It works for symmetric and non symmetric cost;
- Regularity allows to use high order methods;
- It allows to interpolate between different costs and for each $\eta \in [0, 1]$ it returns the solution to the corresponding multi-marginal problems.
- Wasserstein geodesics, Barycenter problem and Martingale transport;

Conclusion

Take-home messages:

- An ODE to characterize entropic multi-marginal optimal transport;
- It works for symmetric and non symmetric cost;
- Regularity allows to use high order methods;
- It allows to interpolate between different costs and for each $\eta \in [0, 1]$ it returns the solution to the corresponding multi-marginal problems.
- Wasserstein geodesics, Barycenter problem and Martingale transport;

Thank You!!