Numerical Methods for Multi-Marginal optimal transport and Density Functional Theory

Luca Nenna

(LMO) Université Paris-Sud

OT Methods for Density Functional Theory, 31/01/2019, Banff

Entropic Optimal Transport

Classical vs Multi-Marginal Optimal Transport

Let us consider two probability measures $\mu, \nu \in \mathcal{P}(X)$ (with $X \subset \mathbb{R}^d$) and a continuous function $c : X \times X \to \mathbb{R}$ then the Monge-Kantorovich formulation $(\mathcal{MK})$ reads as

$$\inf \left\{ \int_{X \times X} c(x,y)\mathbb{P}(x,y)dxdy \mid \quad \mathbb{P} \in \Pi(\mu,\nu) \right\}$$

where $\Pi(\mu,\nu) := \{\mathbb{P} \in \mathcal{P}(X \times X) \mid \quad \pi_{1,\sharp}\mathbb{P} = \mu \quad \pi_{2,\sharp}\mathbb{P} = \nu\}$.
And its extension to the multi-marginal framework

$$\inf \left\{ \int c(x_1, \cdots, x_N)\mathbb{P}(x_1, \cdots, x_N)d\mathbf{x} \mid \mathbb{P} \in \Pi_N(\mu_1, \cdots, \mu_N) \right\} \tag{1}$$

where $\Pi_N(\mu_1, \cdots, \mu_N)$ denotes the set of couplings $\mathbb{P}(x_1, \cdots, x_N)$ having μ_i as marginals.
Remark (Notation): Feel free to take $\mathbb{P}(x_1, \cdots, x_N) = |\Psi|^2$

Classical vs Multi-Marginal Optimal Transport

Let us consider two probability measures $\mu, \nu \in \mathcal{P}(X)$ (with $X \subset \mathbb{R}^d$) and a continuous function $c : X \times X \to \mathbb{R}$ then the Monge-Kantorovich formulation $(\mathcal{MK})$ reads as

$$\inf \left\{ \int_{X \times X} c(x,y)\mathbb{P}(x,y)dxdy | \quad \mathbb{P} \in \Pi(\mu,\nu) \right\}$$

where $\Pi(\mu,\nu) := \{\mathbb{P} \in \mathcal{P}(X \times X) | \quad \pi_{1,\sharp}\mathbb{P} = \mu \quad \pi_{2,\sharp}\mathbb{P} = \nu\}$.
And its extension to the multi-marginal framework

$$\inf \left\{ \int c(x_1,\cdots,x_N)\mathbb{P}(x_1,\cdots,x_N)d\mathbf{x} \mid \mathbb{P} \in \Pi_N(\mu_1,\cdots,\mu_N) \right\} \tag{1}$$

where $\Pi_N(\mu_1,\cdots,\mu_N)$ denotes the set of couplings $\mathbb{P}(x_1,\cdots,x_N)$ having μ_i as marginals.
Remark (Notation): Feel free to take $\mathbb{P}(x_1,\cdots,x_N) = |\Psi|^2$

Some applications

- 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. The transport plan $\mathbb{P}$ matches individuals from each team μ_i minimizing a given cost;
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Paola Gori-Giorgi 2012; C. Cotar, G. Friesecke, and C. Klüppelberg 2013)**). The plan $\mathbb{P}(x_1, \cdots, x_N)$ returns the probability of finding electrons at position $x_1, \cdots, x_N$;
- Incompressible Euler Equations **(Yann Brenier 1989)** : $\mathbb{P}(\omega)$ 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, Di Marino, and L. Nenna 2018)**;
- etc...

Some applications

- 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. The transport plan $\mathbb{P}$ matches individuals from each team μ_i minimizing a given cost;
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Paola Gori-Giorgi 2012; C. Cotar, G. Friesecke, and C. Klüppelberg 2013)**). The plan $\mathbb{P}(x_1, \cdots, x_N)$ returns the probability of finding electrons at position $x_1, \cdots, x_N$;
- Incompressible Euler Equations **(Yann Brenier 1989)** : $\mathbb{P}(\omega)$ 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, Di Marino, and L. Nenna 2018)**;
- etc...

Some applications

- 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. The transport plan $\mathbb{P}$ matches individuals from each team μ_i minimizing a given cost;
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Paola Gori-Giorgi 2012; C. Cotar, G. Friesecke, and C. Klüppelberg 2013)**). The plan $\mathbb{P}(x_1, \cdots, x_N)$ returns the probability of finding electrons at position $x_1, \cdots, x_N$;
- Incompressible Euler Equations **(Yann Brenier 1989)** : $\mathbb{P}(\omega)$ 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, Di Marino, and L. Nenna 2018)**;
- etc...

Some applications

- 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. The transport plan $\mathbb{P}$ matches individuals from each team μ_i minimizing a given cost;
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Paola Gori-Giorgi 2012; C. Cotar, G. Friesecke, and C. Klüppelberg 2013)**). The plan $\mathbb{P}(x_1, \cdots, x_N)$ returns the probability of finding electrons at position $x_1, \cdots, x_N$;
- Incompressible Euler Equations **(Yann Brenier 1989)** : $\mathbb{P}(\omega)$ 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, Di Marino, and L. Nenna 2018)**;
- etc...

Some applications

- 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. The transport plan $\mathbb{P}$ matches individuals from each team μ_i minimizing a given cost;
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Paola Gori-Giorgi 2012; C. Cotar, G. Friesecke, and C. Klüppelberg 2013)**). The plan $\mathbb{P}(x_1, \cdots, x_N)$ returns the probability of finding electrons at position $x_1, \cdots, x_N$;
- Incompressible Euler Equations **(Yann Brenier 1989)** : $\mathbb{P}(\omega)$ 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, Di Marino, and L. Nenna 2018)**;
- etc...

Some applications

- 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. The transport plan $\mathbb{P}$ matches individuals from each team μ_i minimizing a given cost;
- In Density Functional Theory: the electron-electron repulsion (see **(Buttazzo, De Pascale, and Paola Gori-Giorgi 2012; C. Cotar, G. Friesecke, and C. Klüppelberg 2013)**). The plan $\mathbb{P}(x_1, \cdots, x_N)$ returns the probability of finding electrons at position $x_1, \cdots, x_N$;
- Incompressible Euler Equations **(Yann Brenier 1989)** : $\mathbb{P}(\omega)$ 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, Di Marino, and L. Nenna 2018)**;
- etc...

The three universes of Numerical Optimal Transportation

Let's consider the two marginal case then we can have the three following numerical approach to Optimal Transport

- Discrete-2-Discrete: the marginals μ have an atomic form, i.e. $\mu(x) = \sum_i \mu_i \delta_{x_i}$ (and ν as well). **Remarks:**
  - The problem becomes a standard linear programming problem.
  - Works for any kind of cost function.
  - Can be easily generalized to the multi-marginal case.
- Continous-2-Discrete: $\mu = \bar{\mu} dx$ and $\nu(y) = \sum_i \nu_i \delta_{y_i}$. **Remarks:**
  - The semi-discrete approach **(Mérigot 2011)**.
  - Used for generalized euler equations (kind of mmot problem) à la Brenier **(Mérigot and Mirebeau 2016)**.
- Continous-2-Continous $\mu = \bar{\mu} dx$ (and ν too). **Remarks:**
  - The Benamou-Brenier formulation fo Optimal Transport! **(J.-D. Benamou and Y. Brenier 2000)**

The three universes of Numerical Optimal Transportation

Let's consider the two marginal case then we can have the three following numerical approach to Optimal Transport

- Discrete-2-Discrete: the marginals μ have an atomic form, i.e. $\mu(x) = \sum_i \mu_i \delta_{x_i}$ (and ν as well). **Remarks:**
  - The problem becomes a standard linear programming problem.
  - Works for any kind of cost function.
  - Can be easily generalized to the multi-marginal case.
- Continous-2-Discrete: $\mu = \bar{\mu} dx$ and $\nu(y) = \sum_i \nu_i \delta_{y_i}$. **Remarks:**
  - The semi-discrete approach **(Mérigot 2011)**.
  - Used for generalized euler equations (kind of mmot problem) à la Brenier **(Mérigot and Mirebeau 2016)**.
- Continous-2-Continous $\mu = \bar{\mu} dx$ (and ν too). **Remarks**
  - The Benamou-Brenier formulation fo Optimal Transport! **(J.-D. Benamou and Y. Brenier 2000)**

The three universes of Numerical Optimal Transportation

Let's consider the two marginal case then we can have the three following numerical approach to Optimal Transport

- Discrete-2-Discrete: the marginals μ have an atomic form, i.e. $\mu(x) = \sum_i \mu_i \delta_{x_i}$ (and ν as well). **Remarks:**
  - The problem becomes a standard linear programming problem.
  - Works for any kind of cost function.
  - Can be easily generalized to the multi-marginal case.
- Continous-2-Discrete: $\mu = \bar{\mu} dx$ and $\nu(y) = \sum_i \nu_i \delta_{y_i}$. **Remarks:**
  - The semi-discrete approach **(Mérigot 2011)**.
  - Used for generalized euler equations (kind of mmot problem) à la Brenier **(Mérigot and Mirebeau 2016)**.
- Continous-2-Continous $\mu = \bar{\mu} dx$ (and ν too). **Remarks**
  - The Benamou-Brenier formulation fo Optimal Transport! **(J.-D. Benamou and Y. Brenier 2000)**

The three universes of Numerical Optimal Transportation

Let's consider the two marginal case then we can have the three following numerical approach to Optimal Transport

- Discrete-2-Discrete: the marginals μ have an atomic form, i.e. $\mu(x) = \sum_i \mu_i \delta_{x_i}$ (and ν as well). **Remarks:**
  - The problem becomes a standard linear programming problem.
  - Works for any kind of cost function.
  - Can be easily generalized to the multi-marginal case.
- Continous-2-Discrete: $\mu = \bar{\mu} dx$ and $\nu(y) = \sum_i \nu_i \delta_{y_i}$. **Remarks:**
  - The semi-discrete approach **(Mérigot 2011)**.
  - Used for generalized euler equations (kind of mmot problem) à la Brenier **(Mérigot and Mirebeau 2016)**.
- Continous-2-Continous $\mu = \bar{\mu} dx$ (and ν too). **Remarks**
  - The Benamou-Brenier formulation fo Optimal Transport! **(J.-D. Benamou and Y. Brenier 2000)**

The discretized Monge-Kantorovich problem

Let's take $c_{ij} = c(x_i, y_j) \in \mathbb{R}^{M\times M}$ (M are the gridpoints used to discretize X) then the discretized $(\mathcal{MK})$, reads as

$$\min\{\sum_{i,j=1}^{M} c_{ij}\mathbb{P}_{ij} \mid \sum_{j=1}^{M}\mathbb{P}_{ij} = \mu_i \ \forall i,\ \sum_{i=1}^{M}\mathbb{P}_{ij} = \nu_j \ \forall j\} \tag{2}$$

and the dual problem

$$\max\{\sum_{i=1}^{M}\phi_i\mu_i + \sum_{j=1}^{M}\psi_j\nu_j \mid \phi_i + \psi_j \le c_{ij} \ \forall (i,j) \in \{1,\cdots,M\}^2\}. \tag{3}$$

Remarks

- The primal has M^2 unknowns and $M \times 2$ linear constraints.
- The dual has $M \times 2$ unknowns, but M^2 constraints.

The discretized Monge-Kantorovich problem

Let's take $c_{ij} = c(x_i, y_j) \in \mathbb{R}^{M\times M}$ (M are the gridpoints used to discretize X) then the discretized $(\mathcal{MK})$, reads as

$$\min\{\sum_{i,j=1}^{M} c_{ij}\mathbb{P}_{ij} \mid \sum_{j=1}^{M}\mathbb{P}_{ij} = \mu_i\ \forall i,\ \sum_{i=1}^{M}\mathbb{P}_{ij} = \nu_j\ \forall j\} \tag{2}$$

and the dual problem

$$\max\{\sum_{i=1}^{M}\phi_i\mu_i + \sum_{j=1}^{M}\psi_j\nu_j \mid \phi_i + \psi_j \leq c_{ij}\ \forall (i,j) \in \{1,\cdots,M\}^2\}. \tag{3}$$

Remarks

- The primal has M^2 unknowns and $M \times 2$ linear constraints.
- The dual has $M \times 2$ unknowns, but M^2 constraints.

The importance of being sparse

A multi-scale approach to reduce M **(J.-D. Benamou, G. Carlier, and L. Nenna 2016)**

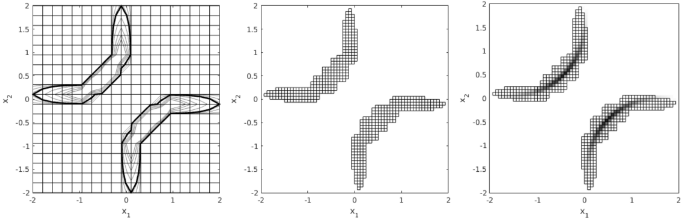


Figure: Support of the optimal $\mathbb{P}$ for 2 marginals and the Coulomb cost

The discretized Monge-Kantorovich problem

Let's take $c_{j_1,\cdots,j_N} = c(x_{j_1},\cdots,x_{j_N}) \in \otimes_1^N \mathbb{R}^M$ (M are the gridpoints used to discretize $\mathbb{R}^d$) then the discretized $(\mathcal{MK}_N)$, reads as

$$\min\{\sum_{(j_1,\cdots,j_N)=1}^{M} c_{j_1,\cdots,j_N}\mathbb{P}_{j_1,\cdots,j_N} \quad | \quad \sum_{j_k,k\neq i}\mathbb{P}_{j_1,\cdots,j_{i-1},j_{i+1},\cdots,j_N} = \mu^i_{j_i}\} \tag{4}$$

and the dual problem

$$\max\{\sum_{i=1}^{N}\sum_{j_i=1}^{M} u^i_{j_i}\mu^i_{j_i} \quad | \quad \sum_{k=1}^{N} u^k_{j_k} \leq c_{j_1,\dots,j_N} \quad \forall (j_1,\cdots,j_N)\in\{1,\cdots,M\}^N\}. \tag{5}$$

Drawbacks

- The primal has M^N unknowns and $M \times N$ linear constraints.
- The dual has $M \times N$ unknowns, but M^N constraints.

The discretized Monge-Kantorovich problem

Let's take $c_{j_1,\cdots,j_N} = c(x_{j_1},\cdots,x_{j_N}) \in \otimes_1^N \mathbb{R}^M$ (M are the gridpoints used to discretize $\mathbb{R}^d$) then the discretized $(\mathcal{MK}_N)$, reads as

$$\min\{ \sum_{(j_1,\cdots,j_N)=1}^{M} c_{j_1,\cdots,j_N}\mathbb{P}_{j_1,\cdots,j_N} \quad | \quad \sum_{j_k,k\neq i} \mathbb{P}_{j_1,\cdots,j_{i-1},j_{i+1},\cdots,j_N} = \mu^i_{j_i}\} \tag{4}$$

and the dual problem

$$\max\{\sum_{i=1}^{N}\sum_{j_i=1}^{M} u^i_{j_i}\mu^i_{j_i} \quad | \quad \sum_{k=1}^{N} u^k_{j_k} \leq c_{j_1,\dots,j_N} \quad \forall (j_1,\cdots,j_N) \in \{1,\cdots,M\}^N\}. \tag{5}$$

Drawbacks

- The primal has M^N unknowns and $M \times N$ linear constraints.
- The dual has $M \times N$ unknowns, but M^N constraints.

The entropic OT problem

We present a numerical method to solve the regularized (**(Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré 2015; M. Cuturi 2013; Galichon and Salanié 2009)**) optimal transport problem (let us consider, for simplicity, 2 marginals)

$$\min_{\mathbb{P}\in\mathcal{C}} \sum_{i,j} c_{ij}\mathbb{P}_{ij} + \begin{cases} \epsilon \sum_{ij} \mathbb{P}_{ij} \log\left(\frac{\mathbb{P}_{ij}}{\mu_i \nu_j}\right) & \mathbb{P} \geq 0 \\ +\infty & otherwise \end{cases} . \quad (6)$$

where C is the matrix associated to the cost, $\mathbb{P}$ is the discrete transport plan and $\mathcal{C}$ is the intersection between $\mathcal{C}_1 = \{\mathbb{P} \mid \sum_j \mathbb{P}_{ij} = \mu_i\}$ and $\mathcal{C}_2 = \{\mathbb{P} \mid \sum_i \mathbb{P}_{ij} = \nu_j\}$.

Remark: Think at ϵ as the temperature, then entropic OT is just OT at positive temperature.

The problem (6) can be re-written as

$$\min_{\mathbb{P}\in\mathcal{C}} \mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) \tag{7}$$

where $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) = \sum_{ij} \mathbb{P}_{ij}\left(\log \frac{\mathbb{P}_{ij}}{\bar{\mathbb{P}}_{ij}}\right)$ $(= \mathrm{KL}(\mathbb{P}|\bar{\mathbb{P}})$ aka the Kullback-Leibler divergence) and $\bar{\mathbb{P}}_{ij} = e^{-\frac{c_{ij}}{\epsilon}}\mu_i\nu_j$.

Remarks:

- **Unique and semi-explicit solution** (we will see it in 2/3 minutes!)
- Problem (7) dates back to Schrödinger, see **(Luca Nenna 2016)** (or better give a look at Christian Léonard's web page).
- $\mathcal{H} \to \mathcal{MK}$ as $\epsilon \to 0$. (see **(Guillaume Carlier, Duval, Gabriel Peyré, and Bernhard Schmitzer 2017; Léonard 2012)**).
- The dual problem is an unconstrained optimization problem.

The problem (6) can be re-written as

$$\min_{\mathbb{P}\in\mathcal{C}} \mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) \tag{7}$$

where $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) = \sum_{ij} \mathbb{P}_{ij}\left(\log \frac{\mathbb{P}_{ij}}{\bar{\mathbb{P}}_{ij}}\right)$ ($= \mathrm{KL}(\mathbb{P}|\bar{\mathbb{P}})$ aka the Kullback-Leibler divergence) and $\bar{\mathbb{P}}_{ij} = e^{-\frac{c_{ij}}{\epsilon}}\mu_i\nu_j$.

Remarks:

- **Unique and semi-explicit solution** (we will see it in 2/3 minutes!)
- Problem (7) dates back to Schrödinger, see **(Luca Nenna 2016)** (or better give a look at Christian Léonard's web page).
- $\mathcal{H} \to \mathcal{MK}$ as $\epsilon \to 0$. (see **(Guillaume Carlier, Duval, Gabriel Peyré, and Bernhard Schmitzer 2017; Léonard 2012)**).
- The dual problem is an unconstrained optimization problem.

The problem (6) can be re-written as

$$\min_{\mathbb{P}\in\mathcal{C}} \mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) \tag{7}$$

where $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) = \sum_{ij} \mathbb{P}_{ij}\left(\log \frac{\mathbb{P}_{ij}}{\bar{\mathbb{P}}_{ij}}\right)$ $(= \mathrm{KL}(\mathbb{P}|\bar{\mathbb{P}})$ aka the Kullback-Leibler divergence) and $\bar{\mathbb{P}}_{ij} = e^{-\frac{c_{ij}}{\epsilon}}\mu_i\nu_j$.

Remarks:

- **Unique and semi-explicit solution** (we will see it in 2/3 minutes!)
- Problem (7) dates back to Schrödinger, see **(Luca Nenna 2016)** (or better give a look at Christian Léonard's web page).
- $\mathcal{H} \to \mathcal{MK}$ as $\epsilon \to 0$. (see **(Guillaume Carlier, Duval, Gabriel Peyré, and Bernhard Schmitzer 2017; Léonard 2012)**).
- The dual problem is an unconstrained optimization problem.

The problem (6) can be re-written as

$$\min_{\mathbb{P}\in\mathcal{C}} \mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) \tag{7}$$

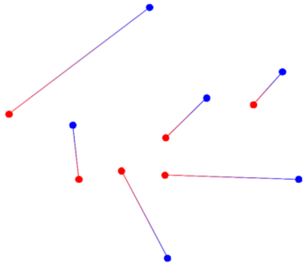
where $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) = \sum_{ij} \mathbb{P}_{ij}\left(\log \frac{\mathbb{P}_{ij}}{\bar{\mathbb{P}}_{ij}}\right)$ $(= \mathrm{KL}(\mathbb{P}|\bar{\mathbb{P}})$ aka the Kullback-Leibler divergence) and $\bar{\mathbb{P}}_{ij} = e^{-\frac{c_{ij}}{\epsilon}}\mu_i\nu_j$.

Remarks:

- **Unique and semi-explicit solution** (we will see it in 2/3 minutes!)
- Problem (7) dates back to Schrödinger, see **(Luca Nenna 2016)** (or better give a look at Christian Léonard's web page).
- $\mathcal{H} \to \mathcal{MK}$ as $\epsilon \to 0$. (see **(Guillaume Carlier, Duval, Gabriel Peyré, and Bernhard Schmitzer 2017; Léonard 2012)**).
- The dual problem is an unconstrained optimization problem.

The "bridge" between quadratic Monge-Kantorovich and Schrödinger

From deterministic to stochastic matching **(Léonard 2012)**

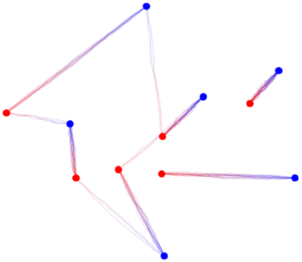


$$\varepsilon = 0$$

Figure: G. Peyre's twitter account

The "bridge" between quadratic Monge-Kantorovich and Schrödinger

From deterministic to stochastic matching **(Léonard 2012)**

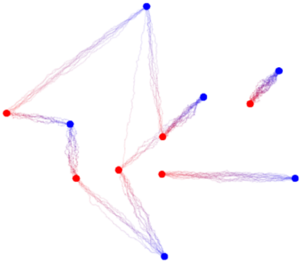


$\varepsilon = .05$

Figure: G. Peyre's twitter account

The "bridge" between quadratic Monge-Kantorovich and Schrödinger

From deterministic to stochastic matching **(Léonard 2012)**

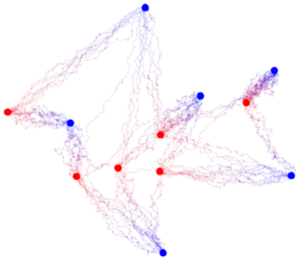


$$\varepsilon = 0.2$$

Figure: G. Peyre's twitter account

The "bridge" between quadratic Monge-Kantorovich and Schrödinger

From deterministic to stochastic matching **(Léonard 2012)**



$$\varepsilon = 1$$

Figure: G. Peyre's twitter account

The Sinkhorn algorithm

Theorem ((Franklin and Lorenz 1989))

The optimal plan $\mathbb{P}^\star$ *has the form* $\mathbb{P}^\star_{ij} = a^\star_i b^\star_j \bar{\mathbb{P}}_{ij}$. *Moreover* $a^\star_i$ *and* $b^\star_j$ *can be uniquely determined (up to a multiplicative constant) as follows*

$$b^\star_j = \frac{\nu_j}{\sum_i a^\star_i \bar{\mathbb{P}}_{ij}}, \; a^\star_i = \frac{\mu_i}{\sum_j b^\star_j \bar{\mathbb{P}}_{ij}}$$

The Sinkhorn algorithm (aka IPFP)

$$b^{n+1}_j = \frac{\nu_j}{\sum_i a^n_i \bar{\mathbb{P}}_{ij}}, \; a^{n+1}_i = \frac{\mu_i}{\sum_j b^{n+1}_j \bar{\mathbb{P}}_{ij}}$$

Theorem ((ibid.))

a^n and b^n converge to $a^\star$ and $b^\star$

Remark: $\phi_i = \epsilon \log(a_i)$ and $\psi_j = \epsilon \log(b_j)$ are the (regularized) Kantorovich potentials.

The Sinkhorn algorithm

Theorem ((Franklin and Lorenz 1989))

The optimal plan $\mathbb{P}^\star$ has the form $\mathbb{P}^\star_{ij} = a^\star_i b^\star_j \bar{\mathbb{P}}_{ij}$. Moreover $a^\star_i$ and $b^\star_j$ can be uniquely determined (up to a multiplicative constant) as follows

$$b^\star_j = \frac{\nu_j}{\sum_i a^\star_i \bar{\mathbb{P}}_{ij}}, \quad a^\star_i = \frac{\mu_i}{\sum_j b^\star_j \bar{\mathbb{P}}_{ij}}$$

The Sinkhorn algorithm (aka IPFP)

$$b^{n+1}_j = \frac{\nu_j}{\sum_i a^n_i \bar{\mathbb{P}}_{ij}}, \quad a^{n+1}_i = \frac{\mu_i}{\sum_j b^{n+1}_j \bar{\mathbb{P}}_{ij}}$$

Theorem ((ibid.))

a^n and b^n converge to $a^\star$ and $b^\star$

Remark: $\phi_i = \epsilon \log(a_i)$ and $\psi_j = \epsilon \log(b_j)$ are the (regularized) Kantorovich potentials.

The Sinkhorn algorithm

Theorem ((Franklin and Lorenz 1989))

The optimal plan $\mathbb{P}^\star$ *has the form* $\mathbb{P}^\star_{ij} = a^\star_i b^\star_j \bar{\mathbb{P}}_{ij}$. *Moreover* $a^\star_i$ *and* $b^\star_j$ *can be uniquely determined (up to a multiplicative constant) as follows*

$$b^\star_j = \frac{\nu_j}{\sum_i a^\star_i \bar{\mathbb{P}}_{ij}}, \; a^\star_i = \frac{\mu_i}{\sum_j b^\star_j \bar{\mathbb{P}}_{ij}}$$

The Sinkhorn algorithm (aka IPFP)

$$b^{n+1}_j = \frac{\nu_j}{\sum_i a^n_i \bar{\mathbb{P}}_{ij}}, \; a^{n+1}_i = \frac{\mu_i}{\sum_j b^{n+1}_j \bar{\mathbb{P}}_{ij}}$$

Theorem ((ibid.))

a^n *and* b^n *converge to* $a^\star$ *and* $b^\star$

Remark: $\phi_i = \epsilon \log(a_i)$ and $\psi_j = \epsilon \log(b_j)$ are the (regularized) Kantorovich potentials.

The Sinkhorn algorithm

Theorem ((Franklin and Lorenz 1989))

The optimal plan $\mathbb{P}^\star$ has the form $\mathbb{P}^\star_{ij} = a^\star_i b^\star_j \bar{\mathbb{P}}_{ij}$. Moreover $a^\star_i$ and $b^\star_j$ can be uniquely determined (up to a multiplicative constant) as follows

$$b^\star_j = \frac{\nu_j}{\sum_i a^\star_i \bar{\mathbb{P}}_{ij}}, \; a^\star_i = \frac{\mu_i}{\sum_j b^\star_j \bar{\mathbb{P}}_{ij}}$$

The Sinkhorn algorithm (aka IPFP)

$$b^{n+1}_j = \frac{\nu_j}{\sum_i a^n_i \bar{\mathbb{P}}_{ij}}, \; a^{n+1}_i = \frac{\mu_i}{\sum_j b^{n+1}_j \bar{\mathbb{P}}_{ij}}$$

Theorem ((ibid.))

a^n and b^n converge to $a^\star$ and $b^\star$

Remark: $\phi_i = \epsilon \log(a_i)$ and $\psi_j = \epsilon \log(b_j)$ are the (regularized) Kantorovich potentials.

Some Remarks

- In **(Franklin and Lorenz 1989)** proved the convergence of Sinkhorn by using the Hilbert metric.
- The entropic regularization spreads the support and this helps to stabilize: it defines a strongly convex program with a unique solution.
- The solution can be obtained through elementary operations (trivially parallelizable).
- The regularized solution $\mathbb{P}^\epsilon$ converges to the solution $\mathbb{P}^{ot}$ of $\mathcal{MK}$ pb. with minimal entropy as $\epsilon \to 0$ (in **(Cominetti and San Martin 1994)** the authors proved that the convergence is exponential).
- The complexity depends on the cost function: with Euler's cost $\mathcal{O}((N-1)M^{2.37})$...still exponential in N for the Coulomb cost **:(** .

Some Remarks

- In **(Franklin and Lorenz 1989)** proved the convergence of Sinkhorn by using the Hilbert metric.
- The entropic regularization spreads the support and this helps to stabilize: it defines a strongly convex program with a unique solution.
- The solution can be obtained through elementary operations (trivially parallelizable).
- The regularized solution $\mathbb{P}^{\epsilon}$ converges to the solution $\mathbb{P}^{ot}$ of $\mathcal{MK}$ pb. with minimal entropy as $\epsilon \to 0$ (in **(Cominetti and San Martin 1994)** the authors proved that the convergence is exponential).
- The complexity depends on the cost function: with Euler's cost $\mathcal{O}((N-1)M^{2.37})$...still exponential in N for the Coulomb cost **:(** .

Some Remarks

- In **(Franklin and Lorenz 1989)** proved the convergence of Sinkhorn by using the Hilbert metric.
- The entropic regularization spreads the support and this helps to stabilize: it defines a strongly convex program with a unique solution.
- The solution can be obtained through elementary operations (trivially parallelizable).
- The regularized solution $\mathbb{P}^\epsilon$ converges to the solution $\mathbb{P}^{ot}$ of $\mathcal{MK}$ pb. with minimal entropy as $\epsilon \to 0$ (in **(Cominetti and San Martin 1994)** the authors proved that the convergence is exponential).
- The complexity depends on the cost function: with Euler's cost $\mathcal{O}((N-1)M^{2.37})$...still exponential in N for the Coulomb cost **:(** .

Some Remarks

- In **(Franklin and Lorenz 1989)** proved the convergence of Sinkhorn by using the Hilbert metric.
- The entropic regularization spreads the support and this helps to stabilize: it defines a strongly convex program with a unique solution.
- The solution can be obtained through elementary operations (trivially parallelizable).
- The regularized solution $\mathbb{P}^\epsilon$ converges to the solution $\mathbb{P}^{ot}$ of $\mathcal{MK}$ pb. with minimal entropy as $\epsilon \to 0$ (in **(Cominetti and San Martin 1994)** the authors proved that the convergence is exponential).
- The complexity depends on the cost function: with Euler's cost $\mathcal{O}((N-1)M^{2.37})$...still exponential in N for the Coulomb cost **:(** .

Some Remarks

- In **(Franklin and Lorenz 1989)** proved the convergence of Sinkhorn by using the Hilbert metric.
- The entropic regularization spreads the support and this helps to stabilize: it defines a strongly convex program with a unique solution.
- The solution can be obtained through elementary operations (trivially parallelizable).
- The regularized solution $\mathbb{P}^{\epsilon}$ converges to the solution $\mathbb{P}^{ot}$ of $\mathcal{MK}$ pb. with minimal entropy as $\epsilon \to 0$ (in **(Cominetti and San Martin 1994)** the authors proved that the convergence is exponential).
- The complexity depends on the cost function: with Euler's cost $\mathcal{O}((N-1)M^{2.37})$...still exponential in N for the Coulomb cost **:(** .

How the regularization works: from spread to deterministic plan (quadratic cost)

Take the quadratic cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

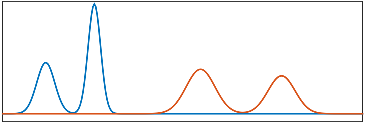


Figure: Marginals μ and ν

Figure: $\epsilon = 60/N$

How the regularization works: from spread to deterministic plan (quadratic cost)

Take the quadratic cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

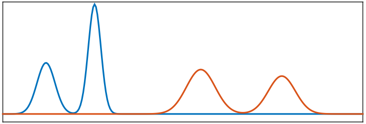


Figure: Marginals μ and ν

Figure: $\epsilon = 40/N$

How the regularization works: from spread to deterministic plan (quadratic cost)

Take the quadratic cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

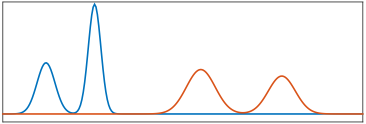


Figure: Marginals μ and ν

Figure: $\epsilon = 20/N$

How the regularization works: from spread to deterministic plan (quadratic cost)

Take the quadratic cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

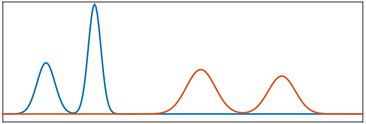


Figure: Marginals μ and ν

Figure: $\epsilon = 10/N$

How the regularization works: from spread to deterministic plan (quadratic cost)

Take the quadratic cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

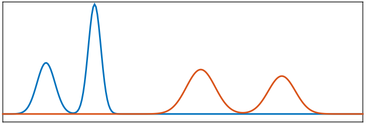


Figure: Marginals μ and ν

Figure: $\epsilon = 6/N$

How the regularization works: from spread to deterministic plan (quadratic cost)

Take the quadratic cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

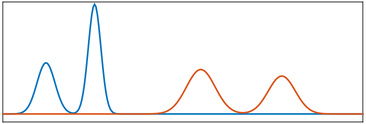


Figure: Marginals μ and ν

Figure: $\epsilon = 4/N$

The extension to the Multi-Marginal problem

The entropic multi-marginal problem becomes

$$\min_{\mathbb{P}\in\mathcal{C}} \mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) \tag{8}$$

where $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) = \sum_{i,j,k} \mathbb{P}_{ijk}(\log \frac{\mathbb{P}_{ijk}}{\bar{\mathbb{P}}_{ijk}} - 1)$ is the relative entropy, and $\mathcal{C} = \bigcap_{i=1}^{3} \mathcal{C}_i$ (i.e. $\mathcal{C}_1 = \{\mathbb{P} \;\; | \;\; \sum_{j,k} \mathbb{P}_{ijk} = \mu_i^1\}$).

The optimal plan $\mathbb{P}^\star$ becomes $\mathbb{P}^\star_{ijk} = a_i^\star b_j^\star c_k^\star \bar{\mathbb{P}}_{ijk}$ $a_i^\star$, $b_j^\star$ and $c_k^\star$ can be determined by the marginal constraints.

$$b_j^\star = \frac{\mu_j^2}{\sum_{ik} a_i^\star c_k^\star \bar{\mathbb{P}}_{ijk}} \quad \Rightarrow \quad b_j^{n+1} = \frac{\mu_j^2}{\sum_{ik} a_i^n c_k^n \bar{\mathbb{P}}_{ijk}}$$

$$c_k^\star = \frac{\mu_k^3}{\sum_{ij} a_i^\star b_j^\star \bar{\mathbb{P}}_{ijk}} \quad \Rightarrow \quad c_k^{n+1} = \frac{\mu_k^3}{\sum_{ij} a_i^n b_j^{n+1} \bar{\mathbb{P}}_{ijk}}$$

$$a_i^\star = \frac{\mu_i^1}{\sum_{jk} b_j^\star c_k^\star \bar{\mathbb{P}}_{ijk}} \quad \Rightarrow \quad a_i^{n+1} = \frac{\mu_i^1}{\sum_{jk} b_j^{n+1} c_k^{n+1} \bar{\mathbb{P}}_{ijk}}$$

The extension to the Multi-Marginal problem

The entropic multi-marginal problem becomes

$$\min_{\mathbb{P}\in\mathcal{C}} \mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) \tag{8}$$

where $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) = \sum_{i,j,k} \mathbb{P}_{ijk}(\log \frac{\mathbb{P}_{ijk}}{\bar{\mathbb{P}}_{ijk}} - 1)$ is the relative entropy, and $\mathcal{C} = \bigcap_{i=1}^{3} \mathcal{C}_i$ (i.e. $\mathcal{C}_1 = \{\mathbb{P} \;\mid\; \sum_{j,k} \mathbb{P}_{ijk} = \mu_i^1\}$).

The optimal plan $\mathbb{P}^\star$ becomes $\mathbb{P}^\star_{ijk} = a_i^\star b_j^\star c_k^\star \bar{\mathbb{P}}_{ijk}$ $a_i^\star$, $b_j^\star$ and $c_k^\star$ can be determined by the marginal constraints.

$$b_j^\star = \frac{\mu_j^2}{\sum_{ik} a_i^\star c_k^\star \bar{\mathbb{P}}_{ijk}} \quad \Rightarrow \quad b_j^{n+1} = \frac{\mu_j^2}{\sum_{ik} a_i^n c_k^n \bar{\mathbb{P}}_{ijk}}$$

$$c_k^\star = \frac{\mu_k^3}{\sum_{ij} a_i^\star b_j^\star \bar{\mathbb{P}}_{ijk}} \quad \Rightarrow \quad c_k^{n+1} = \frac{\mu_k^3}{\sum_{ij} a_i^n b_j^{n+1} \bar{\mathbb{P}}_{ijk}}$$

$$a_i^\star = \frac{\mu_i^1}{\sum_{jk} b_j^\star c_k^\star \bar{\mathbb{P}}_{ijk}} \quad \Rightarrow \quad a_i^{n+1} = \frac{\mu_i^1}{\sum_{jk} b_j^{n+1} c_k^{n+1} \bar{\mathbb{P}}_{ijk}}$$

The extension to the Multi-Marginal problem

The entropic multi-marginal problem becomes

$$\min_{\mathbb{P}\in\mathcal{C}} \mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) \tag{8}$$

where $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}) = \sum_{i,j,k} \mathbb{P}_{ijk}(\log \frac{\mathbb{P}_{ijk}}{\bar{\mathbb{P}}_{ijk}} - 1)$ is the relative entropy, and $\mathcal{C} = \bigcap_{i=1}^{3} \mathcal{C}_i$ (i.e. $\mathcal{C}_1 = \{\mathbb{P} \;\;|\;\; \sum_{j,k} \mathbb{P}_{ijk} = \mu_i^1\}$).

The optimal plan $\mathbb{P}^\star$ becomes $\mathbb{P}^\star_{ijk} = a_i^\star b_j^\star c_k^\star \bar{\mathbb{P}}_{ijk}$ $a_i^\star$, $b_j^\star$ and $c_k^\star$ can be determined by the marginal constraints.

$$b_j^\star = \frac{\mu_j^2}{\sum_{ik} a_i^\star c_k^\star \bar{\mathbb{P}}_{ijk}} \quad\Rightarrow\quad b_j^{n+1} = \frac{\mu_j^2}{\sum_{ik} a_i^n c_k^n \bar{\mathbb{P}}_{ijk}}$$

$$c_k^\star = \frac{\mu_k^3}{\sum_{ij} a_i^\star b_j^\star \bar{\mathbb{P}}_{ijk}} \quad\Rightarrow\quad c_k^{n+1} = \frac{\mu_k^3}{\sum_{ij} a_i^n b_j^{n+1} \bar{\mathbb{P}}_{ijk}}$$

$$a_i^\star = \frac{\mu_i^1}{\sum_{jk} b_j^\star c_k^\star \bar{\mathbb{P}}_{ijk}} \quad\Rightarrow\quad a_i^{n+1} = \frac{\mu_i^1}{\sum_{jk} b_j^{n+1} c_k^{n+1} \bar{\mathbb{P}}_{ijk}}$$

Sinkhornizing the world!!

- Wasserstein Barycenter **(Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré 2015)**;
- Matching for teams **(Luca Nenna 2016)**;
- Optimal transport with capacity constraint **(Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré 2015)**;
- Partial Optimal Transport **(Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré 2015; Chizat, G. Peyré, B. Schmitzer, and Vialard 2016)**;
- Multi-Marginal Optimal Transport **(Luca Nenna 2016; J.-D. Benamou, G. Carlier, and L. Nenna 2016; Jean-David Benamou, Guillaume Carlier, and Luca Nenna 2018; Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré 2015)**;
- Wasserstein Gradient Flows (JKO) **(Gabriel Peyré 2015)**;
- Unbalanced Optimal Transport **(Chizat, G. Peyré, B. Schmitzer, and Vialard 2016)**;
- Cournot-Nash equilibria **(Blanchet, Guillaume Carlier, and Luca Nenna 2017)**
- Mean Field Games **(J.-D. Benamou, G. Carlier, Di Marino, and L. Nenna 2018)**;
- And more is coming...

MMOT with Coulomb cost

The Levy-Lieb functional

Consider the Levy-Lieb functional $F_{LL}[\rho]$

$$F_{LL}[\rho] = \min_{\Psi \to \rho} \epsilon T[\Psi] + V_{ee}[\Psi] \tag{9}$$

Remark (super rough!!!): Let's take $\mathbb{P} = |\Psi|^2$, then $|\nabla \Psi|^2 = |\nabla \sqrt{\mathbb{P}}|^2 = \frac{1}{4} \frac{|\nabla \mathbb{P}|^2}{\mathbb{P}}$ and the kinetic energy can be re-written as

$$T[\Psi] = \int_{\mathbb{R}^{dN}} \frac{1}{4} \frac{|\nabla \mathbb{P}|^2}{\mathbb{P}} dx_1 \cdots dx_N.$$

Then we have **(Bindini and De Pascale 2017; Codina Cotar, Gero Friesecke, and Claudia Klüppelberg 2018; Lewin 2018)**...

Semiclassical limit

$\lim_{\epsilon \to 0} F_{LL}[\rho] = \mathcal{MK}[\rho]$

The Levy-Lieb functional

Consider the Levy-Lieb functional $F_{LL}[\rho]$

$$F_{LL}[\rho] = \min_{\Psi \to \rho} \epsilon T[\Psi] + V_{ee}[\Psi] \tag{9}$$

Remark (super rough!!!): Let's take $\mathbb{P} = |\Psi|^2$, then $|\nabla\Psi|^2 = |\nabla\sqrt{\mathbb{P}}|^2 = \frac{1}{4}\frac{|\nabla\mathbb{P}|^2}{\mathbb{P}}$ and the kinetic energy can be re-written as

$$T[\Psi] = \int_{\mathbb{R}^{dN}} \frac{1}{4}\frac{|\nabla\mathbb{P}|^2}{\mathbb{P}} dx_1 \cdots dx_N.$$

Then we have **(Bindini and De Pascale 2017; Codina Cotar, Gero Friesecke, and Claudia Klüppelberg 2018; Lewin 2018)**...

Semiclassical limit

$\lim_{\epsilon\to 0} F_{LL}[\rho] = \mathcal{MK}[\rho]$

The Levy-Lieb functional

Consider the Levy-Lieb functional $F_{LL}[\rho]$

$$F_{LL}[\rho] = \min_{\Psi \to \rho} \epsilon T[\Psi] + V_{ee}[\Psi] \tag{9}$$

Remark (super rough!!!): Let's take $\mathbb{P} = |\Psi|^2$, then $|\nabla \Psi|^2 = |\nabla \sqrt{\mathbb{P}}|^2 = \frac{1}{4}\frac{|\nabla \mathbb{P}|^2}{\mathbb{P}}$ and the kinetic energy can be re-written as

$$T[\Psi] = \int_{\mathbb{R}^{dN}} \frac{1}{4}\frac{|\nabla \mathbb{P}|^2}{\mathbb{P}} dx_1 \cdots dx_N.$$

Then we have **(Bindini and De Pascale 2017; Codina Cotar, Gero Friesecke, and Claudia Klüppelberg 2018; Lewin 2018)**...

Semiclassical limit

$\lim_{\epsilon \to 0} F_{LL}[\rho] = \mathcal{MK}[\rho]$

The entropic inequality

One can prove the following inequality

The Entropic Inequality (Seidl, Di Marino, Gerolin, L. Nenna, Giesbertz, and P. Gori-Giorgi 2017)

$$\min_{\mathbb{P}\to\rho}\int_{\mathbb{R}^{dN}}\epsilon\frac{1}{4}\frac{|\nabla\mathbb{P}|^2}{\mathbb{P}}+\sum_{i<j}\frac{1}{|x_i-x_j|}\mathbb{P}\geq\min_{\mathbb{P}\to\rho}\int_{\mathbb{R}^{dN}}\epsilon C\mathbb{P}\log(\mathbb{P})+\sum_{i<j}\frac{1}{|x_i-x_j|}\mathbb{P}=\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}}). \tag{10}$$

where $\int\frac{1}{4}\frac{|\nabla\mathbb{P}|^2}{\mathbb{P}}\geq C\int\mathbb{P}\log(\mathbb{P})$ is the log-sobolev inequality (or Fisher information) and the entropic functional $\mathcal{H}(\mathbb{P}|\bar{\mathbb{P}})$ corresponds to minimize the Kullback-Leibler distance between $\mathbb{P}$ and $\bar{\mathbb{P}}=e^{-\sum_{i<j}\frac{1}{|x_i-x_j|}\frac{1}{C\epsilon}}$.

The limit as $\epsilon \to 0$

Take the Coulomb cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

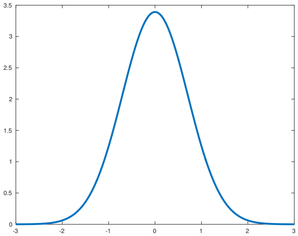


Figure: Marginals ρ (and ρ)

Figure: $\epsilon = 10$

The limit as $\epsilon \to 0$

Take the Coulomb cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

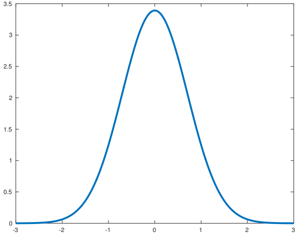


Figure: Marginals ρ (and ρ)

Figure: $\epsilon = 5$

The limit as $\epsilon \to 0$

Take the Coulomb cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

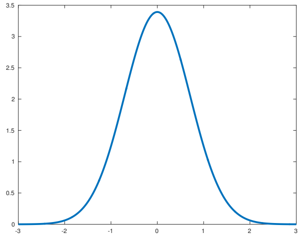


Figure: Marginals ρ (and ρ)

Figure: $\epsilon = 1$

The limit as $\epsilon \to 0$

Take the Coulomb cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

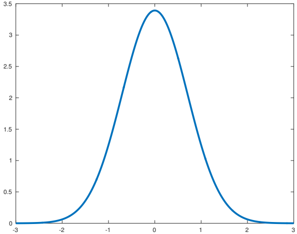


Figure: Marginals ρ (and ρ)

Figure: $\epsilon = 0.1$

The limit as $\epsilon \to 0$

Take the Coulomb cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

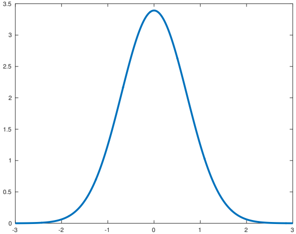


Figure: Marginals ρ (and ρ)

Figure: $\epsilon = 0.01$

The limit as $\epsilon \to 0$

Take the Coulomb cost and solve the regularized problem. Then as $\epsilon \to 0$ ($N = 512$), we have

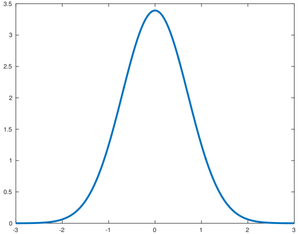
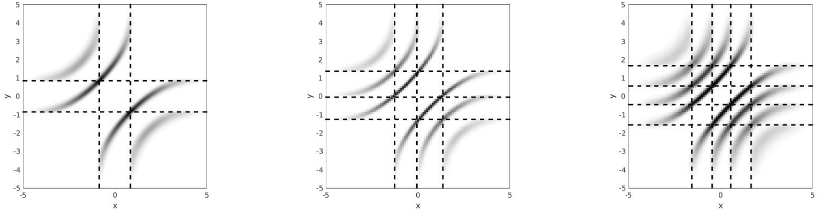


Figure: Marginals ρ (and ρ)

Figure: $\epsilon = 0.002$

Some simulations for $N = 3, 4, 5$ in 1D

We take the density $\rho(x) = \frac{N}{10}(1 + \cos(\frac{\pi}{5}x))$ and...



$N = 3$ $N = 4$ $N = 5$

Figure: *Support of the projected plan* $\pi_{12}(\mathbb{P})$

SGS vs Entropic: the uniform density on the ball ($N = 3$)

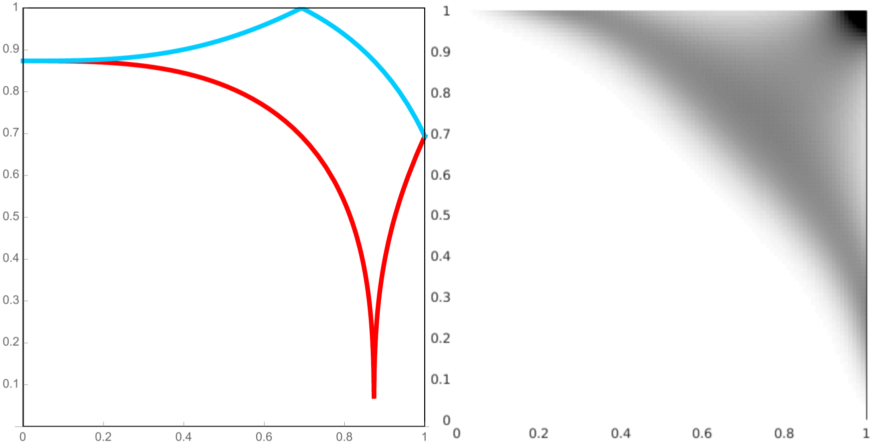


Figure: SGS maps (left) $\mathcal{MK}_{SGS} = 2.32682$ and entropic plan (right) $\mathcal{MK}_{\epsilon} = 2.31721$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...

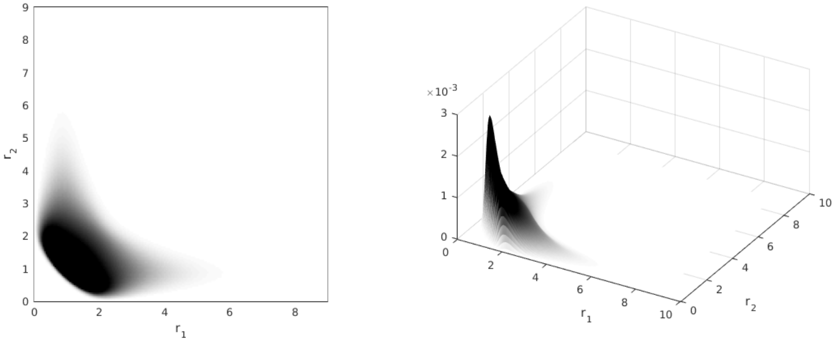


Figure: $\alpha = 0$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...

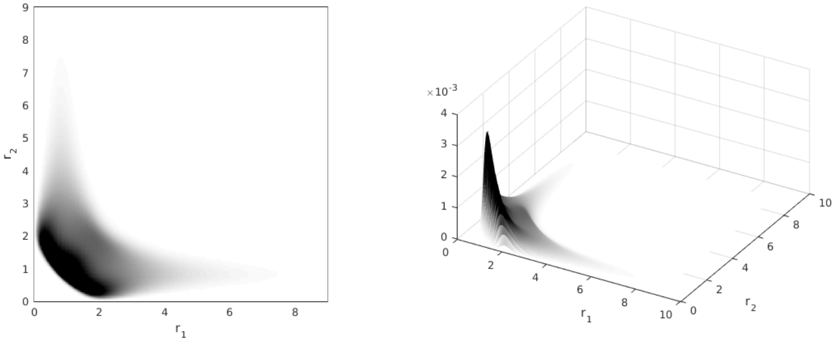


Figure: $\alpha = 0.1429$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...

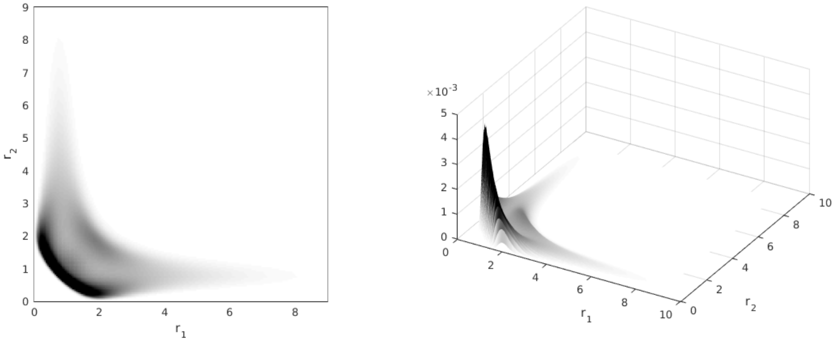


Figure: $\alpha = 0.2857$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...

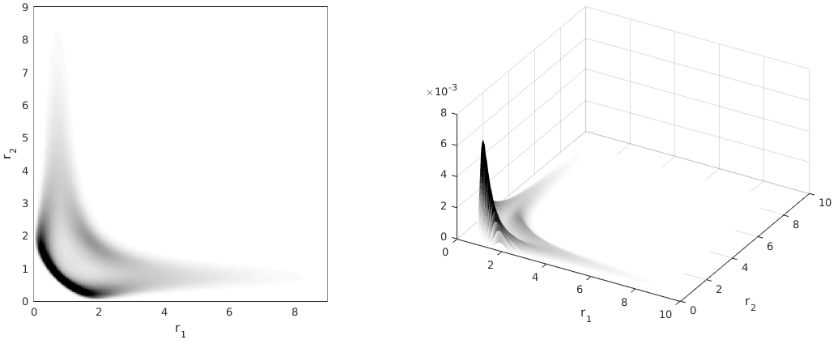


Figure: $\alpha = 0.4286$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...

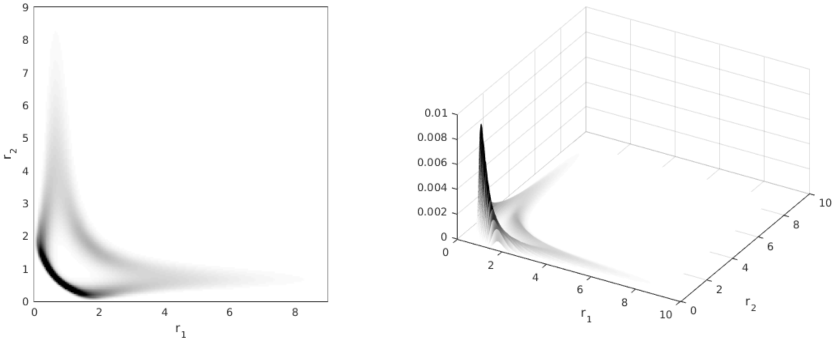


Figure: $\alpha = 0.5714$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...

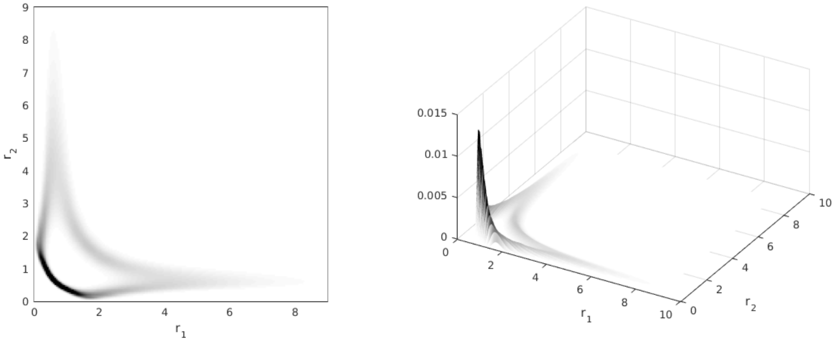
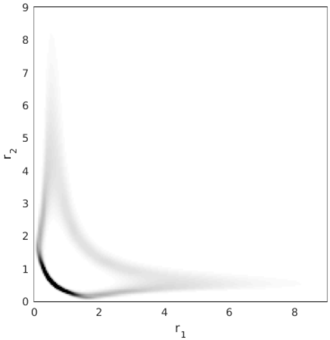


Figure: $\alpha = 0.7143$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...



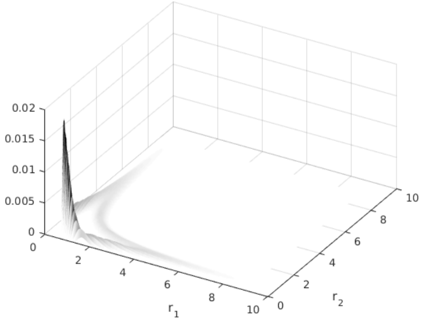


Figure: $\alpha = 0.8571$

The transition from spread to deterministic plans for $N = 3$ and $d = 3$

Take $\rho_\alpha(r) = \alpha\rho_{Li}(r) + (1-\alpha)\rho_{exp}(r)$ and $\alpha \in [0,1]$ then...

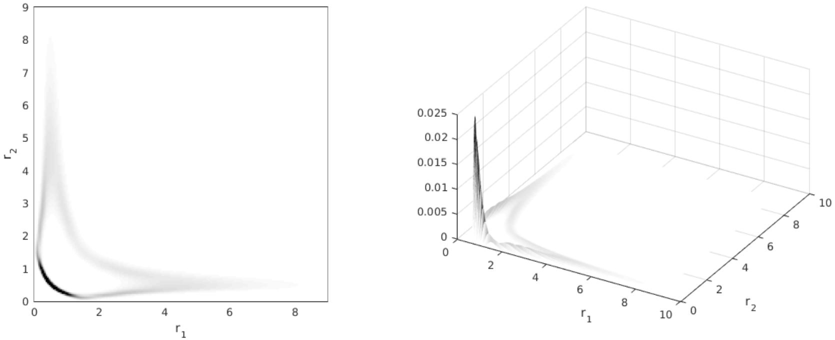


Figure: $\alpha = 1$

Advertising

If you are interested in OT, Entropic regularization and more:

- My web page (just google me) or contact me **luca.nenna@math.u-psud.fr**;
- Mokaplan team **https://team.inria.fr/mokaplan/**;

Some references:

Benamou, J.-D., G. Carlier, & L. Nenna (2016). "A Numerical Method to solve Multi-Marginal Optimal Transport Problems with Coulomb Cost". In: *Splitting Methods in Communication, Imaging, Science, and Engineering*. Springer International Publishing, pp. 577–601.

Benamou, Jean-David, Guillaume Carlier, Marco Cuturi, Luca Nenna, & Gabriel Peyré (2015). "Iterative Bregman projections for regularized transportation problems". In: *SIAM J. Sci. Comput.* 37.2, A1111–A1138. ISSN: 1064-8275. DOI: 10.1137/141000439. URL: http://dx.doi.org/10.1137/141000439.

Nenna, Luca (2016). "Numerical methods for multi-marginal optimal transportation". PhD thesis. PSL Research University.

Peyré, Gabriel & Marco Cuturi (2017). *Computational optimal transport*. Tech. rep.

Thank You!!