Multimarginal optimal transportation: the one dimensional symmetric case

Simone Di Marino
(in collaboration with Maria Colombo and Luigi De Pascale)

March 11, 2014

Introduction

Plan of the talk:

Introduction

Plan of the talk:

- Description of the multi marginal problem and physical motivation;

Introduction

Plan of the talk:

- Description of the multi marginal problem and physical motivation;
- Classical results with 2 marginals in dimension 1;

Introduction

Plan of the talk:

- Description of the multi marginal problem and physical motivation;
- Classical results with 2 marginals in dimension 1;
- Symmetric multimarginal case in dimension 1: existence of an optimal map and uniqueness of the symmetric optimal plan.

General minimization problem

We are interested in the following minimum problem:

$$\min_{f \in L^2((\mathbb{R}^d)^n;\mathbb{C}), \|f\|_2=1} \int_{(\mathbb{R}^d)^n} (c|\nabla f|^2 + V_{ee}|f|^2 + V_{ext}|f|^2)\,\mathrm{d}x,$$

General minimization problem

We are interested in the following minimum problem:

$$\min_{f\in L^2((\mathbb{R}^d)^n;\mathbb{C}),\|f\|_2=1} \int_{(\mathbb{R}^d)^n} (c|\nabla f|^2 + V_{ee}|f|^2 + V_{ext}|f|^2)\,\mathrm{d}x,$$

where

- V_{ee} is the Coulombian interaction potential between the electrons:

$$V_{ee}(x_1,\ldots,x_n) = \sum_{1\leq i<j\leq n} \frac{1}{|x_i - x_j|};$$

- V_{ext} is an external potential (due to the nucleus), which is the same for every electron: $V_{ext}(x_1,\ldots,x_n) = V(x_1) + \cdots + V(x_n)$;

Hohenberg e Kohn formulation(HK)

The wave function f gives us the density ρ of the (identical) electrons, that is, the marginals of the measure $|f|^2 dx$ are always ρ.

Hohenberg e Kohn formulation(HK)

The wave function f gives us the density ρ of the (identical) electrons, that is, the marginals of the measure $|f|^2 dx$ are always ρ.
In order to solve the previous problem, the idea is to fix the density ρ and minimize only the (kinetic + interaction) part under the density constraint. Then one tries to solve the main problem as

$$\min_{\rho \in \mathcal{P}(\mathbb{R}^d)} \left\{ F(\rho) + n \int V(x) \, \mathrm{d}\rho \right\},$$

where

$$F(\rho) = \min_{\rho \leftarrow f} \int_{(\mathbb{R}^d)^n} (c|\nabla f|^2 + V_{ee}|f|^2) \, \mathrm{d}x,$$

Hohenberg e Kohn formulation(HK)

The wave function f gives us the density ρ of the (identical) electrons, that is, the marginals of the measure $|f|^2 dx$ are always ρ.
In order to solve the previous problem, the idea is to fix the density ρ and minimize only the (kinetic + interaction) part under the density constraint. Then one tries to solve the main problem as

$$\min_{\rho \in \mathcal{P}(\mathbb{R}^d)} \left\{ F(\rho) + n \int V(x) \, \mathrm{d}\rho \right\},$$

where

$$F(\rho) = \min_{\rho \leftarrow f} \int_{(\mathbb{R}^d)^n} (c|\nabla f|^2 + V_{ee}|f|^2) \, \mathrm{d}x,$$

If now we take the semiclassical limit $c \to 0$, we have convergence of F to another minimum problem (Cotar - Friesecke - Klüppelberg). We are left is our multimarginal optimal transport problem.

Hohenberg e Kohn formulation(HK)

The wave function f gives us the density ρ of the (identical) electrons, that is, the marginals of the measure $|f|^2 dx$ are always ρ.
In order to solve the previous problem, the idea is to fix the density ρ and minimize only the (kinetic + interaction) part under the density constraint. Then one tries to solve the main problem as

$$\min_{\rho \in \mathcal{P}(\mathbb{R}^d)} \left\{ F(\rho) + n \int V(x) \, \mathrm{d}\rho \right\},$$

where

$$F(\rho) = \min_{\rho \leftarrow f} \int_{(\mathbb{R}^d)^n} (c|\nabla f|^2 + V_{ee}|f|^2) \, \mathrm{d}x,$$

If now we take the semiclassical limit $c \to 0$, we have convergence of F to another minimum problem (Cotar - Friesecke - Klüppelberg). We are left is our multimarginal optimal transport problem.
Another way is to see it as the strictly correlated regime, in which we neglect the kinetic part (Gori Giorgi).

Mathematical formulation

We are interested in the problem

$$\min_{\pi \in \Gamma(\mu_1, \ldots, \mu_n)} \int_{X^n} c(x_1, \ldots, x_n) \, \mathrm{d}\pi.$$

In our case $X = \mathbb{R}^d$, $\mu_1 = \ldots = \mu_n = \rho$ and c is the sum of the repulsive Coulombian potentials.

Mathematical formulation

We are interested in the problem

$$\min_{\pi \in \Gamma(\mu_1, \ldots, \mu_n)} \int_{X^n} c(x_1, \ldots, x_n) \, \mathrm{d}\pi.$$

In our case $X = \mathbb{R}^d$, $\mu_1 = \ldots = \mu_n = \rho$ and c is the sum of the repulsive Coulombian potentials.

- When $n = 2$ we obtain the Kantorovich formulation of optimal transport;

Mathematical formulation

We are interested in the problem

$$\min_{\pi \in \Gamma(\mu_1, \dots, \mu_n)} \int_{X^n} c(x_1, \dots, x_n) \, \mathrm{d}\pi.$$

In our case $X = \mathbb{R}^d$, $\mu_1 = \ldots = \mu_n = \rho$ and c is the sum of the repulsive Coulombian potentials.

- When $n = 2$ we obtain the Kantorovich formulation of optimal transport;
- this isn't exactly a trasport problem but more a coupling problem;

Mathematical formulation

We are interested in the problem

$$\min_{\pi \in \Gamma(\mu_1, \ldots, \mu_n)} \int_{X^n} c(x_1, \ldots, x_n) \, \mathrm{d}\pi.$$

In our case $X = \mathbb{R}^d$, $\mu_1 = \ldots = \mu_n = \rho$ and c is the sum of the repulsive Coulombian potentials.

- When $n = 2$ we obtain the Kantorovich formulation of optimal transport;
- this isn't exactly a trasport problem but more a coupling problem;
- symmetries: impossibility for uniqueness.

Multimarginal optimal transport

Notations:

- given $\sigma \in S_n$ define $\sigma : X^n \to X^n$ as
$$\sigma : (x_1, \dots, x_n) \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(n)});$$
- $\Gamma_{sym}(\rho)$ is the set of probabilities π which have all marginals equal to ρ and such that $\sigma_\sharp \pi = \pi$ for all $\sigma \in S_n$. The natural projection into symmetric plans is
$$\pi^S = \frac{1}{n!} \sum_{\sigma \in S_n} \sigma_\sharp \pi;$$
- $T_{sym}(\rho)$ is the set of Borel maps $T : X \to X$ such that $T_\sharp \rho = \rho$ e $T^{(n)}(x) = x$ for ρ-almost every x.

Multimarginal optimal transport

Notations:

- given $\sigma \in S_n$ define $\sigma : X^n \to X^n$ as
$$\sigma : (x_1, \dots, x_n) \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(n)});$$
- $\Gamma_{sym}(\rho)$ is the set of probabilities π which have all marginals equal to ρ and such that $\sigma_\sharp \pi = \pi$ for all $\sigma \in S_n$. The natural projection into symmetric plans is
$$\pi^S = \frac{1}{n!} \sum_{\sigma \in S_n} \sigma_\sharp \pi;$$
- $T_{sym}(\rho)$ is the set of Borel maps $T : X \to X$ such that $T_\sharp \rho = \rho$ e $T^{(n)}(x) = x$ for ρ-almost every x.

From now on we'll call (K_{sym}) e (M_{sym}) the following two problems:

$$\min_{\pi \in \Gamma_{sym}(\rho)} \int_{X^n} c(x_1, \dots, x_n) \, \mathrm{d}\pi$$

$$\inf_{T \in T_{sym}(\rho)} \int_X c(x, T(x), T(T(x)), \dots, T^{(n-1)}(x)) \, \mathrm{d}\rho$$

Existence of the map

Conjecture

There exists an optimal symmetric map, in particular (M_{sym}) *is a minimum. Furthermore this minimum is equal to* (K_{sym}).

Existence of the map

Conjecture

There exists an optimal symmetric map, in particular (M_{sym}) is a minimum. Furthermore this minimum is equal to (K_{sym}).

The goal of this studies is to investigate the equality $(K_{sym}) = (M_{sym})$, the presence of an optimal map and eventually the characterization of optimal (symmetric) plans.

Known results

- (Colombo - D.M.) In the symmetric case is true in general that $(K_{sym}) = (M_{sym})$, in every complete and separable metric space, with a symmetric l.s.c. cost, continuous in its finiteness domain, when ρ is without atoms;

Known results

- (Colombo - D.M.) In the symmetric case is true in general that $(K_{sym}) = (M_{sym})$, in every complete and separable metric space, with a symmetric l.s.c. cost, continuous in its finiteness domain, when ρ is without atoms;
- (Pass, Pass - Kim) in the non-symmetric case, if a modifed twist condition on the cost holds true and ρ is absolutely continuous with respect to Lebesgue measure, then there exists a unique optimal plan, which is induced by an optimal map (as in the two marginal case);

Known results

- (Colombo - D.M.) In the symmetric case is true in general that $(K_{sym}) = (M_{sym})$, in every complete and separable metric space, with a symmetric l.s.c. cost, continuous in its finiteness domain, when ρ is without atoms;
- (Pass, Pass - Kim) in the non-symmetric case, if a modifed twist condition on the cost holds true and ρ is absolutely continuous with respect to Lebesgue measure, then there exists a unique optimal plan, which is induced by an optimal map (as in the two marginal case);
- (Cotar - Friesecke - Klüppelberg) In the case $N = 2$ there is existence of the map, in any dimension (radial case);

Known results

- (Colombo - D.M.) In the symmetric case is true in general that $(K_{sym}) = (M_{sym})$, in every complete and separable metric space, with a symmetric l.s.c. cost, continuous in its finiteness domain, when ρ is without atoms;
- (Pass, Pass - Kim) in the non-symmetric case, if a modifed twist condition on the cost holds true and ρ is absolutely continuous with respect to Lebesgue measure, then there exists a unique optimal plan, which is induced by an optimal map (as in the two marginal case);
- (Cotar - Friesecke - Klüppelberg) In the case $N = 2$ there is existence of the map, in any dimension (radial case);

Theorem (Buttazzo - De Pascale - Gori Giorgi)

There exists a minimizer of (K_{sym}). Furthermore, the following duality formula holds true

$$(K_{sym}) = n \sup \left\{ \int_X \phi \, \mathrm{d}\rho \ : \ \phi(x_1) + \ldots + \phi(x_n) \leq c(x_1, \ldots, x_n) \right\}$$

Let $X = \mathbb{R}$, and let $c(x_1, \ldots, x_n) = \sum_{1 \le i < j \le n} \phi(|x_i - x_j|)$, where ϕ is a convex and decreasing function on $\mathbb{R}^+$.

Let $X = \mathbb{R}$, and let $c(x_1, \ldots, x_n) = \sum_{1 \le i < j \le n} \phi(|x_i - x_j|)$, where ϕ is a convex and decreasing function on $\mathbb{R}^+$.

Theorem

Let ρ be a diffuse probability measure on $\mathbb{R}$ (such that $(K) < \infty$). Let $-\infty = d_0 < d_1 < \ldots < d_N = +\infty$ be such that

$$\rho([d_i, d_{i+1}]) = 1/N \qquad \forall i = 0, \ldots, N-1. \tag{1}$$

Let $T : \mathbb{R} \to \mathbb{R}$ be the unique function (up to ρ-null sets) that is increasing on every interval $[d_i, d_{i+1}]$, $i = 0, ..., N-1$, and such that

$$T_\sharp 1_{[d_i, d_{i+1}]}\rho = 1_{[d_{i+1}, d_{i+2}]}\rho \quad \forall i = 0, \ldots, N-1. \tag{2}$$

Then T is an admissible map for (M_{sym}) and

$$(K) = \int_{\mathbb{R}} c(x, T(x), T^2(x), \ldots, T^{(N-1)}(x))\, d\rho. \tag{3}$$

Moreover there exists a unique symmetric optimal plan, that is the symmetrization of the one induced by T.

c-monotonicity

Definition (c-monotonicity)

A set $A \subset X \times X$ is c-monotone if for every two points $(x_1, y_1), (x_2, y_2) \in A$ we have

$$c(x_1, y_1) + c(x_2, y_2) \leq c(x_1, y_2) + c(x_2, y_1)$$

c-monotonicity

Definition (c-monotonicity)

A set $A \subset X \times X$ is c-monotone if for every two points $(x_1, y_1), (x_2, y_2) \in A$ we have

$$c(x_1, y_1) + c(x_2, y_2) \leq c(x_1, y_2) + c(x_2, y_1)$$

Theorem (Gangbo)

Let c be a nonnegative l.s.c. cost, continuous in its finiteness domain; then if the optimal cost is finite, the support of every optimal plan is c-monotone.

c-monotonicity

Definition (c-monotonicity)

A set $A \subset X \times X$ is c-monotone if for every two points $(x_1, y_1), (x_2, y_2) \in A$ we have

$$c(x_1, y_1) + c(x_2, y_2) \leq c(x_1, y_2) + c(x_2, y_1)$$

Theorem (Gangbo)

Let c be a nonnegative l.s.c. cost, continuous in its finiteness domain; then if the optimal cost is finite, the support of every optimal plan is c-monotone.

Proof.

Consider neighborhoods of the points where inequality fails and rearrange the plan. □

Geometric characterization in $\mathbb{R}$, convex cost

In dimension 1, in the case $c(x,y)=\phi(x-y)$, with ϕ strictly convex function we have: if A is c-monotone, the for every couple $(x_1,y_1),(x_2,y_2)\in A$, if $x_1<x_2$ then $y_1\leq y_2$.

Geometric characterization in $\mathbb{R}$, convex cost

In dimension 1, in the case $c(x,y) = \phi(x-y)$, with ϕ strictly convex function we have: if A is c-monotone, the for every couple $(x_1, y_1), (x_2, y_2) \in A$, if $x_1 < x_2$ then $y_1 \leq y_2$.

Proof.

Suppose $y_1 > y_2$. Then we have that $x_1 - y_1 < x_2 - y_1 < x_2 - y_2$, so there exists $t \in (0,1)$ such that

$$x_2 - y_1 = t(x_1 - y_1) + (1-t)(x_2 - y_2)$$

Geometric characterization in $\mathbb{R}$, convex cost

In dimension 1, in the case $c(x,y) = \phi(x-y)$, with ϕ strictly convex function we have: if A is c-monotone, the for every couple $(x_1,y_1),(x_2,y_2) \in A$, if $x_1 < x_2$ then $y_1 \leq y_2$.

Proof.

Suppose $y_1 > y_2$. Then we have that $x_1 - y_1 < x_2 - y_1 < x_2 - y_2$, so there exists $t \in (0,1)$ such that

$$x_2 - y_1 = t(x_1 - y_1) + (1-t)(x_2 - y_2)$$

$$x_1 - y_2 = (1-t)(x_1 - y_1) + t(x_2 - y_2)$$

Geometric characterization in $\mathbb{R}$, convex cost

In dimension 1, in the case $c(x, y) = \phi(x - y)$, with ϕ strictly convex function we have: if A is c-monotone, the for every couple $(x_1, y_1), (x_2, y_2) \in A$, if $x_1 < x_2$ then $y_1 \leq y_2$.

Proof.

Suppose $y_1 > y_2$. Then we have that $x_1 - y_1 < x_2 - y_1 < x_2 - y_2$, so there exists $t \in (0, 1)$ such that

$$\phi(x_2 - y_1) < t\phi(x_1 - y_1) + (1 - t)\phi(x_2 - y_2)$$

$$\phi(x_1 - y_2) < (1 - t)\phi(x_1 - y_1) + t\phi(x_2 - y_2)$$

□

Geometric characterization in $\mathbb{R}$, convex cost

In dimension 1, in the case $c(x,y) = \phi(x-y)$, with ϕ strictly convex function we have: if A is c-monotone, the for every couple $(x_1,y_1),(x_2,y_2) \in A$, if $x_1 < x_2$ then $y_1 \leq y_2$.

Proof.

Suppose $y_1 > y_2$. Then we have that $x_1 - y_1 < x_2 - y_1 < x_2 - y_2$, so there exists $t \in (0,1)$ such that

$$\phi(x_2 - y_1) < t\phi(x_1 - y_1) + (1-t)\phi(x_2 - y_2)$$

$$\phi(x_1 - y_2) < (1-t)\phi(x_1 - y_1) + t\phi(x_2 - y_2)$$

$$\phi(x_2 - y_1) + \phi(x_1 - x_2) < \phi(x_1 - y_1) + \phi(x_2 - y_2)$$

□

Corollary

The support of an optimal plan is "monotone".If ρ_1 and ρ_2 are diffuse then an optimal plan is induced by a map.

multimarginal c-monotonicity

We can adapt the definition to the multi marginal case. Given $P \subseteq \{1, \dots, n\}$ and two points $x, y \in X^n$ we define the P-mixing as

$$(P(x,y))_i = \begin{cases} x_i & \text{if } i \in P \\ y_i & \text{otherwise.} \end{cases} \qquad (P^c(x,y))_i = \begin{cases} y_i & \text{if } i \in P \\ x_i & \text{otherwise.} \end{cases} .$$

Definition (c-monotonicity)

A set $A \subset X^n$ is c-monotone if for every $x, y \in A$ and every $P \subseteq \{1, \dots, n\}$ one has

$$c(x) + c(y) \leq c(P(x,y)) + c(P^c(x,y))$$

multimarginal c-monotonicity

We can adapt the definition to the multi marginal case. Given $P \subseteq \{1, \dots, n\}$ and two points $x, y \in X^n$ we define the P-mixing as

$$(P(x,y))_i = \begin{cases} x_i & \text{if } i \in P \\ y_i & \text{otherwise.} \end{cases} \qquad (P^c(x,y))_i = \begin{cases} y_i & \text{if } i \in P \\ x_i & \text{otherwise.} \end{cases} .$$

Definition (c-monotonicity)

A set $A \subset X^n$ is c-monotone if for every $x, y \in A$ and every $P \subseteq \{1, \dots, n\}$ one has

$$c(x) + c(y) \leq c(P(x,y)) + c(P^c(x,y))$$

Theorem (Pass)

Let c be a nonnegative l.s.c. cost, continuous in its finiteness domain; then the support of an optimal plan with finite cost is necessarily c-monotone.

The proof uses Gangbo theorem applied to the spaces $X^{|P|}$ e $X^{|P^c|}$.

symmetric c-monotonicity

If we add symmetry we can define c-monotonicity in the following equivalent way. We first define the "cumulative coordinate" of two points $x, y \in X^n$ as

$$C(x,y) = \sum_{i=1}^{n} \delta_{x_i} + \delta_{y_i}$$

Definition (c-monotonicity)

A symmetric set $A \subset X^n$ is c-monotone if for every points $x, y \in A$ and every points x', y' such that $C(x,y) = C(x',y')$ we have that

$$c(x) + c(y) \leq c(x') + c(y')$$

From now on we'll have $X = \mathbb{R}$ and $c(x) = \sum_{i<j} \phi(|x_i - x_j|)$. The crucial geometrical characterization is:

From now on we'll have $X = \mathbb{R}$ and $c(x) = \sum_{i<j} \phi(|x_i - x_j|)$. The crucial geometrical characterization is:

Lemma

Given a set T consisting of $2n$ coordinates counted with multiplicities, i.e., there exist $t_1 \leq t_2 \leq \cdots \leq t_{2n}$ such that $T = \sum \delta_{t_i}$, let x, y be points in X^n that minimize the problem

$$\min\{c(x') + c(y') \ : \ C(x', y') = T\}.$$

Then, up to swap x and y and up to re-arrange the coordinates increasingly, we have that $x_i = t_{2i}$ and $y_i = t_{2i-1}$.

From now on we'll have $X = \mathbb{R}$ and $c(x) = \sum_{i<j} \phi(|x_i - x_j|)$. The crucial geometrical characterization is:

Lemma

Given a set T consisting of $2n$ coordinates counted with multiplicities, i.e., there exist $t_1 \leq t_2 \leq \cdots \leq t_{2n}$ such that $T = \sum \delta_{t_i}$, let x, y be points in X^n that minimize the problem

$$\min\{c(x') + c(y') \ : \ C(x', y') = T\}.$$

Then, up to swap x and y and up to re-arrange the coordinates increasingly, we have that $x_i = t_{2i}$ and $y_i = t_{2i-1}$.

Equivalently, x, y are minimizers iff, defining x^*, y^* the increasing rearrangement of the coordinates, one has

$$x_1^* \leq y_1^* \leq x_2^* \leq \cdots x_n^* \leq y_n^*;$$

in this case the points are *well ordered*.

How to conclude

Given the lemma, we can conclude the proof of the theorem:

- prove that the support of an optimal plan doesn't intersect $x_i = x_j$;
- reduce the analysis to the zone $x_1 < x_2 \cdots < x_n = O$;
- consider the numbers

$$d_i^+ = \max\{x_i : x \in O \cap \mathrm{spt}\pi\} \qquad d_i^- = \min\{x_i : x \in O \cap \mathrm{spt}\pi\},$$

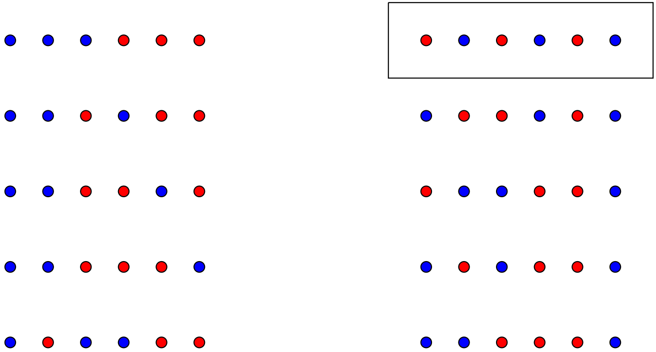
  and prove $d_i^+ \leq d_{i+1}^-$ and consequently $\rho([d_i^+, d_{i+1}^+]) = 1/n$;
- conclude as in the 2 dimensional case, with monotone maps.

Case with 3 marginals: by hands

We have 6 points on the real line (that represent T). In what way I can group in two groups of three points such that the interaction potential is minimal? $\binom{6}{3}/2 = 10$ cases.

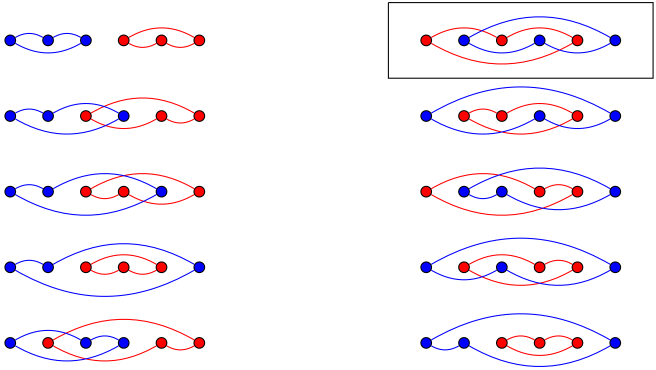
Case with 3 marginals: by hands

We have 6 points on the real line (that represent T). In what way I can group in two groups of three points such that the interaction potential is minimal? $\binom{6}{3}/2 = 10$ cases.



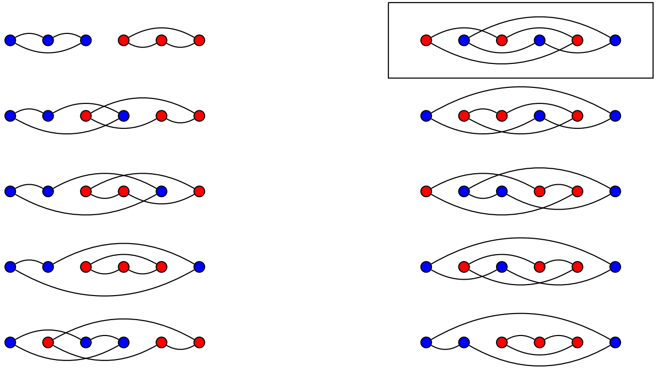
Case with 3 marginals: by hands

We have 6 points on the real line (that represent T). In what way I can group in two groups of three points such that the interaction potential is minimal? $\binom{6}{3}/2 = 10$ cases.



Case with 3 marginals: by hands

We have 6 points on the real line (that represent T). In what way I can group in two groups of three points such that the interaction potential is minimal? $\binom{6}{3}/2 = 10$ cases.

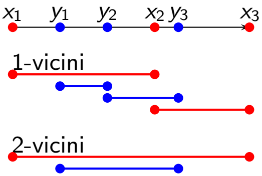


n-marginals symmetric case

Given $2n$ points on the real line, what is the best way of separate them in two group of n points such that the energy interaction is minimal?

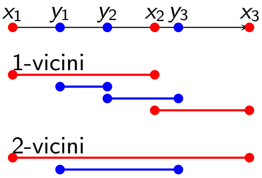
n-marginals symmetric case

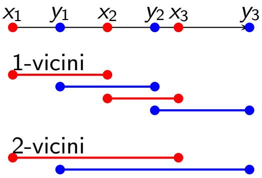
Given $2n$ points on the real line, what is the best way of separate them in two group of n points such that the energy interaction is minimal? Let us consider k-neighborhoods:



n-marginals symmetric case

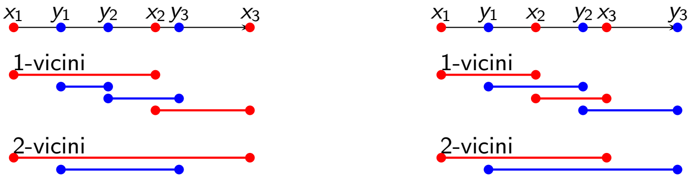
Given $2n$ points on the real line, what is the best way of separate them in two group of n points such that the energy interaction is minimal? Let us consider k-neighborhoods:





n-marginals symmetric case

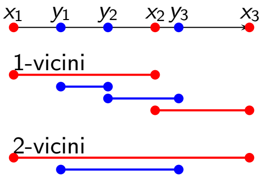
Given $2n$ points on the real line, what is the best way of separate them in two group of n points such that the energy interaction is minimal? Let us consider k-neighborhoods:

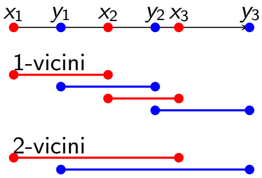


- Step 1 (convexity of ϕ): use the one-dimensional result for 2 marginals on k-neighbors (with fixed k);

n-marginals symmetric case

Given $2n$ points on the real line, what is the best way of separate them in two group of n points such that the energy interaction is minimal? Let us consider k-neighborhoods:





- Step 1 (convexity of ϕ): use the one-dimensional result for 2 marginals on k-neighbors (with fixed k);
- Step 2 (monotonicity of ϕ): stretch in the right way the segments to get the "well ordered" situation.

Future development

- one-dimensional symmetric case with a cost that is concave or concave-convex;

Future development

- one-dimensional symmetric case with a cost that is concave or concave-convex;
- symmetric case for $X = \mathbb{R}^2$, with radial marginal and Coulombian cost;

Future development

- one-dimensional symmetric case with a cost that is concave or concave-convex;
- symmetric case for $X = \mathbb{R}^2$, with radial marginal and Coulombian cost;
- sufficiency of multimarginal c-monotonicity (other possible definitions?).

Thanks for the attention