Étudier l'espace des diagrammes de persistance grâce au transport optimal

GdT Analyse Topologique des Données
18 juin 2019

Théo Lacombe

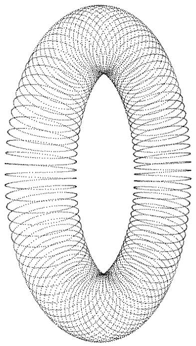
DataShape - Inria Saclay

theo.lacombe@inria.fr

The TDA pipeline: persistent homology

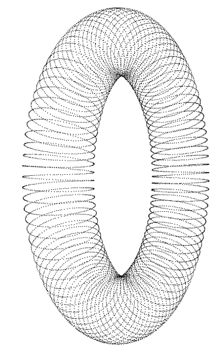
A (very) concise summary:

Input: some (complex) object



The TDA pipeline: persistent homology

A (very) concise summary:



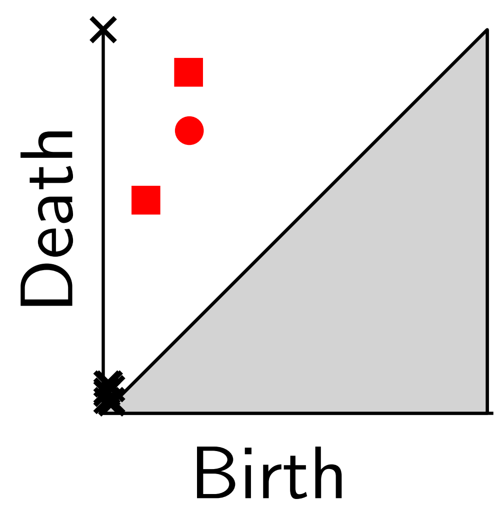
Input: some (complex) object

$\downarrow$ filtration f

Sequence of increasing topo. spaces

$\downarrow$ persistent homology theory

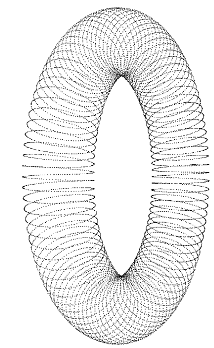
Persistence diagram



- Multiset of points
  $\{(x_1, n_1) \ldots (x_i, n_i) \ldots\}$

The TDA pipeline: persistent homology

A (very) concise summary:



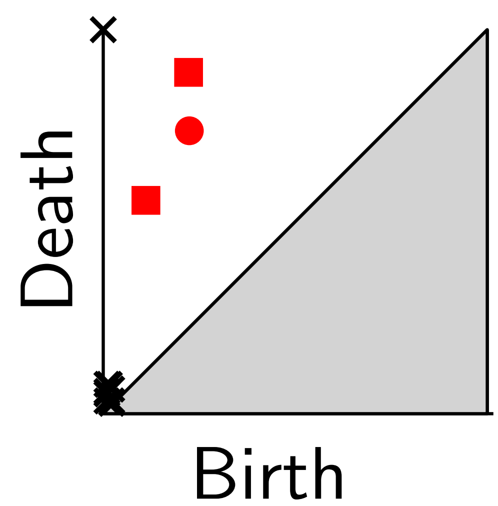
Input: some (complex) object

↓ filtration f

Sequence of increasing topo. spaces

↓ persistent homology theory

Persistence diagram



- Radon measures (locally finite Borel measures)
  $\{(x_1, n_1) \ldots (x_i, n_i) \ldots\} \leftrightarrow \sum_i n_i \delta_{x_i}$

The TDA pipeline: persistent homology

A (very) concise summary:

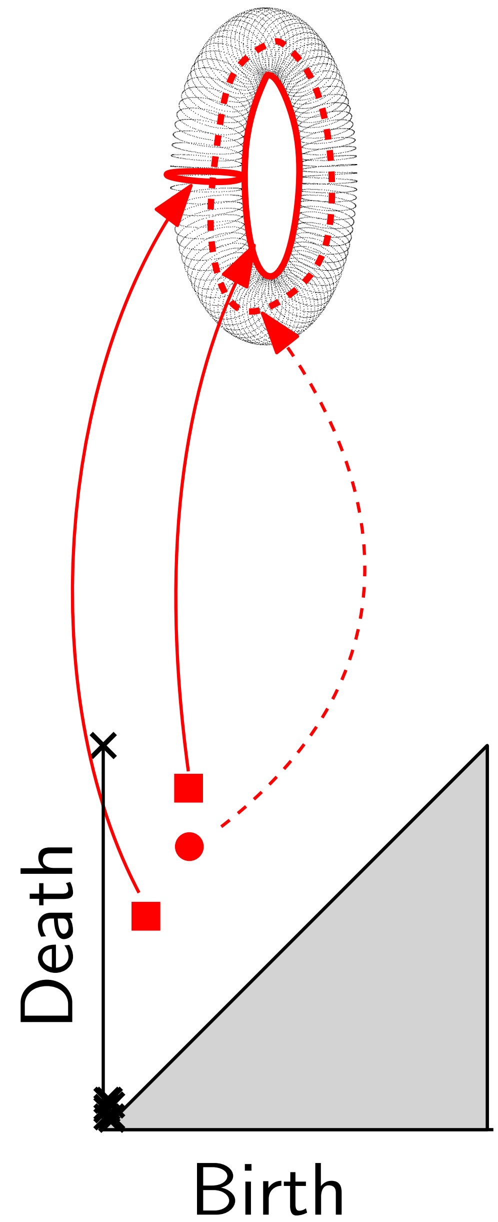
Input: some (complex) object

$\downarrow$ filtration f

Sequence of increasing topo. spaces

$\downarrow$ persistent homology theory

Persistence diagram



- Radon measures (locally finite Borel measures)
  $\{(x_1, n_1) \ldots (x_i, n_i) \ldots\} \leftrightarrow \sum_i n_i \delta_{x_i}$

The space of persistence diagrams $(\mathcal{D}^p, d_p)$

Persistence diagrams are:

- Interpretable

The space of persistence diagrams $(\mathcal{D}^p, d_p)$

Persistence diagrams are:

- Interpretable
- Comparable, using matching-like metrics

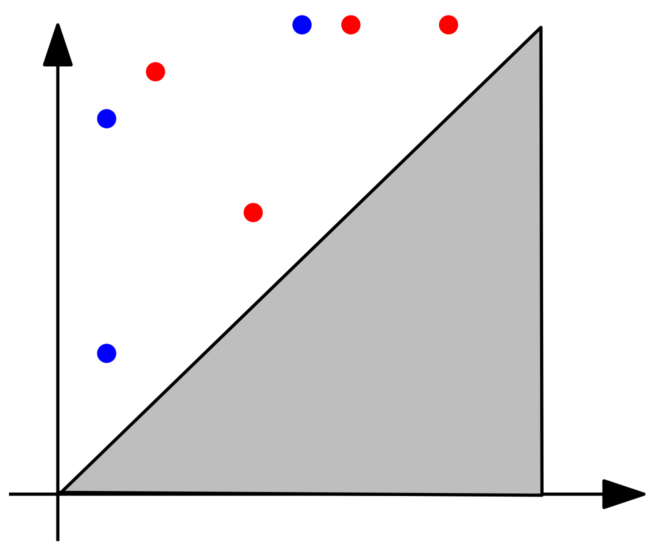
$$d_p(a,b) = \left(\inf_{\zeta} \sum_{(x,y)\in\zeta} \|x-y\|^p + \sum_{u\notin\zeta} \|u-s(u)\|^p \right)^{1/p}$$

partial matching

unmatched points

matched points

proj on diag



The space of persistence diagrams $(\mathcal{D}^p, d_p)$

Persistence diagrams are:

- Interpretable
- Comparable, using matching-like metrics

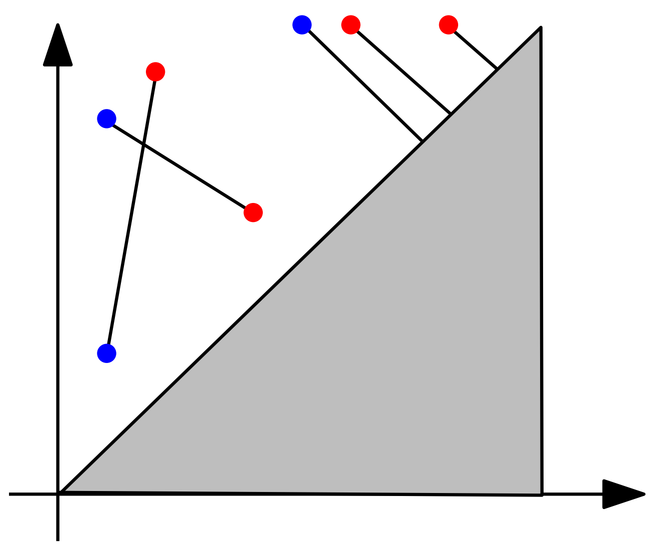
$$d_p(a,b) = \left(\inf_{\zeta} \sum_{(x,y)\in\zeta} \|x - y\|^p + \sum_{u \notin \zeta} \|u - s(u)\|^p \right)^{1/p}$$

partial matching

unmatched points

proj on diag

matched points



The space of persistence diagrams $(\mathcal{D}^p, d_p)$

Persistence diagrams are:

- Interpretable
- Comparable, using matching-like metrics

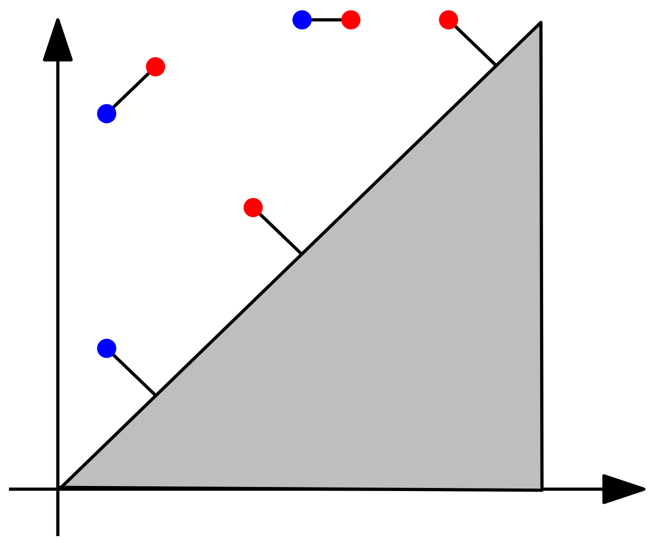
$$d_p(a,b) = \left(\inf_{\zeta} \sum_{(x,y)\in\zeta} \|x-y\|^p + \sum_{u\notin\zeta} \|u-s(u)\|^p\right)^{1/p}$$

partial matching

unmatched points

proj on diag

matched points



The space of persistence diagrams $(\mathcal{D}^p, d_p)$

Persistence diagrams are:

- Interpretable
- Comparable, using matching-like metrics

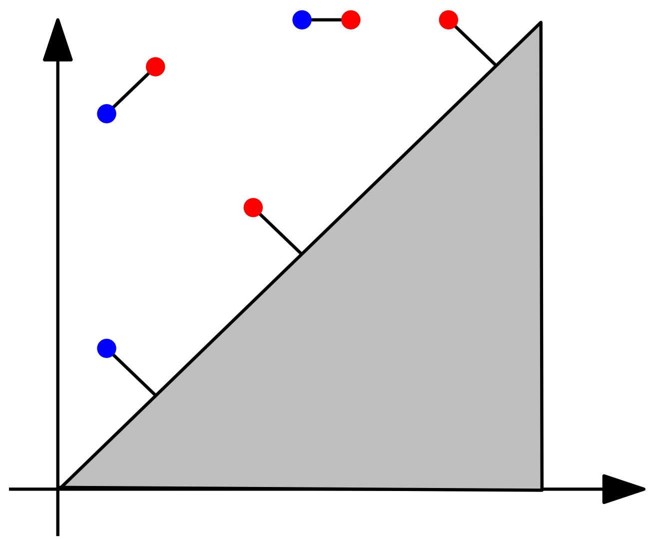
$$d_p(a, b) = \left(\inf_{\zeta} \sum_{(x,y)\in\zeta} \|x - y\|^p + \sum_{u\notin\zeta} \|u - s(u)\|^p \right)^{1/p}$$

partial matching

unmatched points

matched points

proj on diag

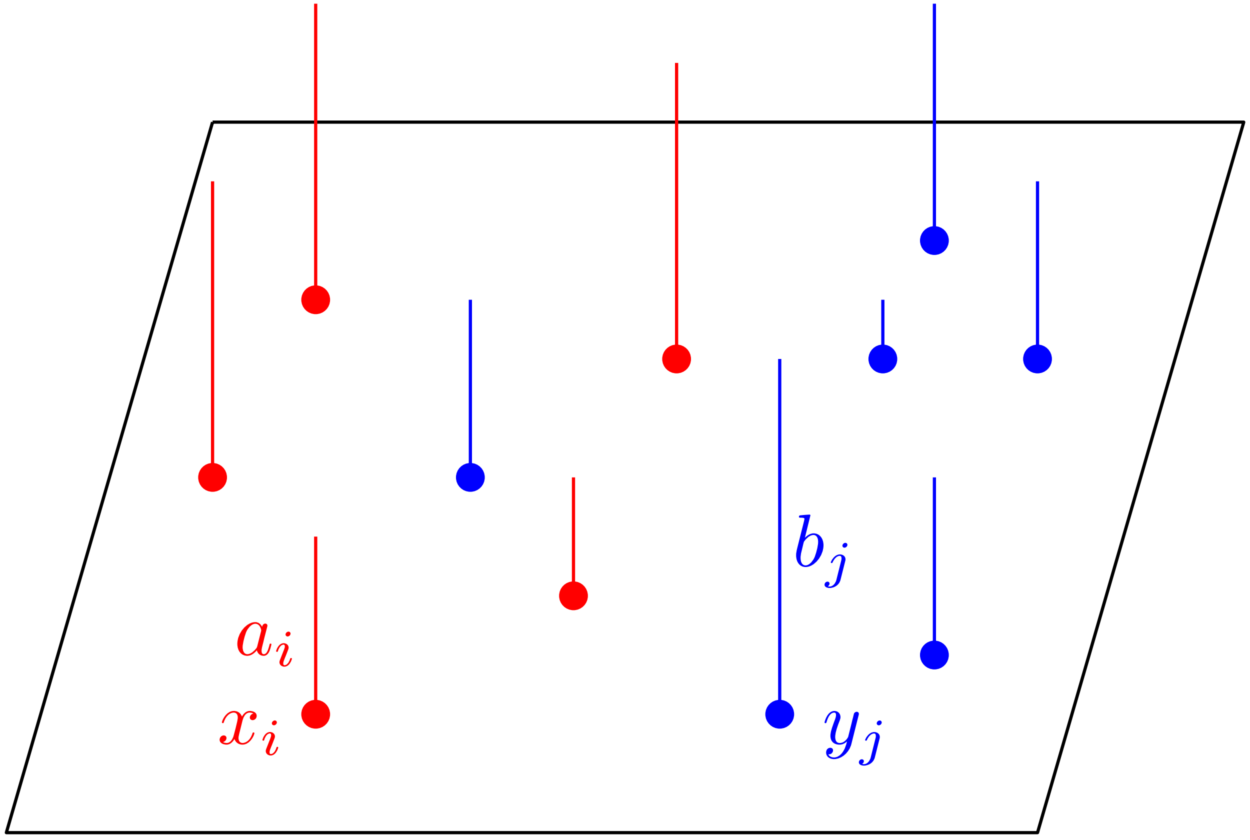


- Theoretically motivated
- Stable wrt input data

Optimal Transport - Generalities

Discrete formulation : μ and ν two probability measures

$$\mu = \sum_i a_i \delta_{x_i} \qquad \nu = \sum_j b_j \delta_{y_j}$$

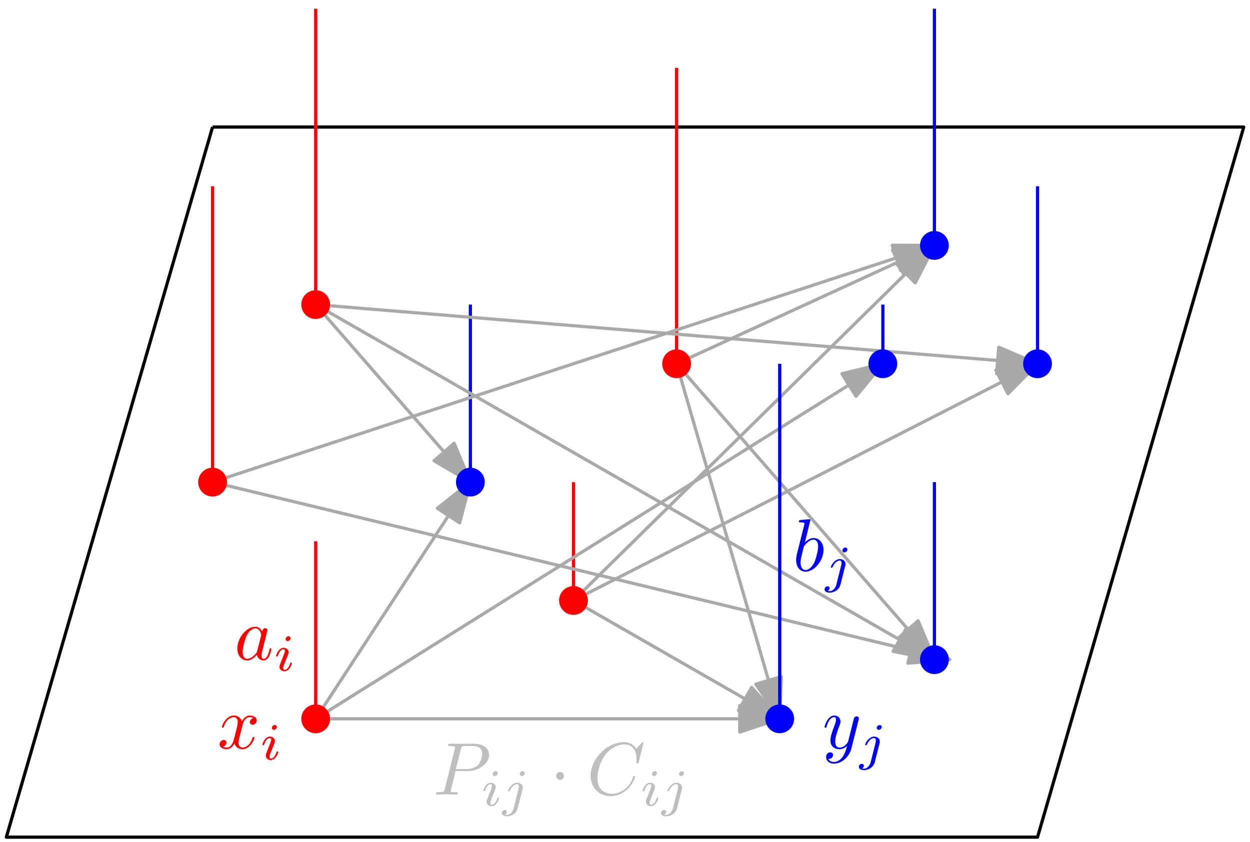


Optimal Transport - Generalities

Discrete formulation : μ and ν two probability measures

$$\mu = \sum_i a_i \delta_{x_i} \qquad \nu = \sum_j b_j \delta_{y_j}$$

(P_{ij}) transport plan between μ and ν



$$\sum_j P_{ij} = a_i$$

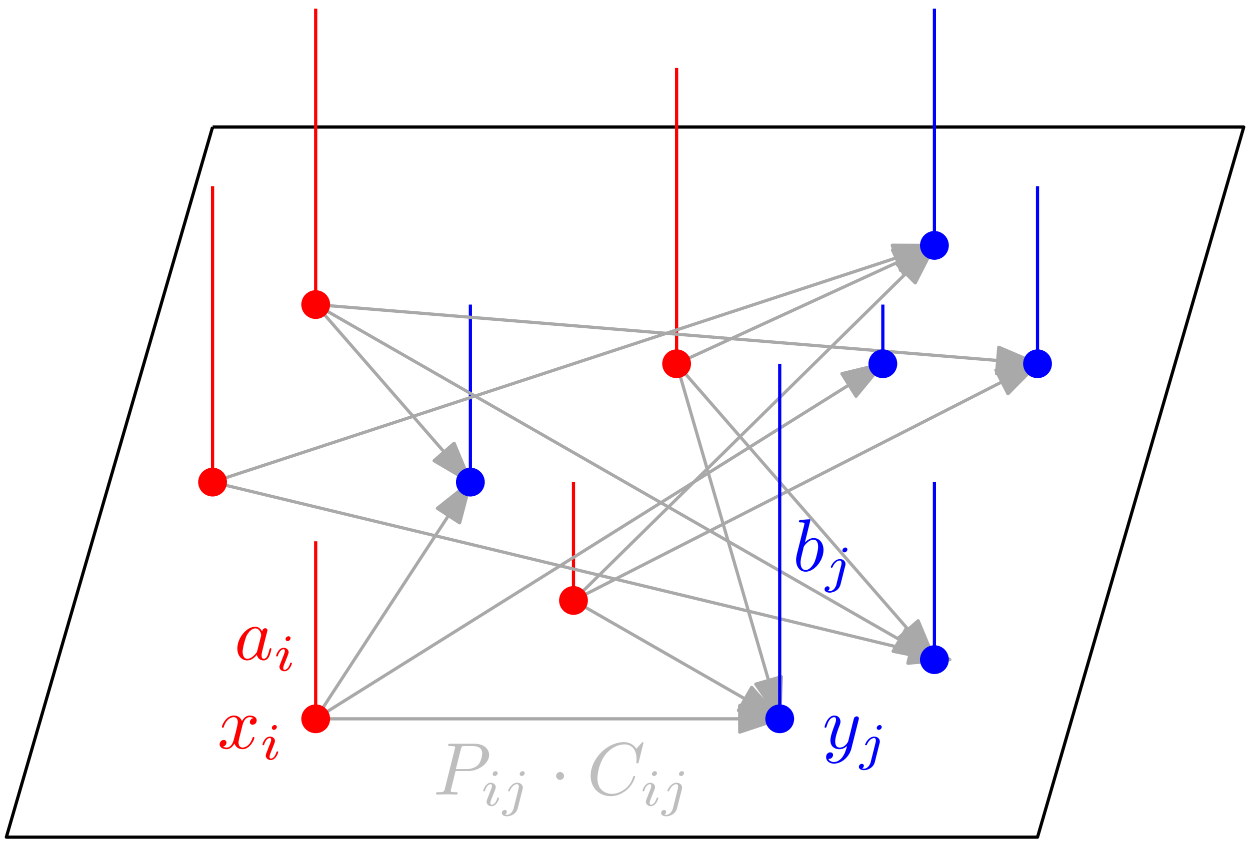
$$\sum_i P_{ij} = b_j$$

Optimal Transport - Generalities

Discrete formulation : μ and ν two probability measures

$$\mu = \sum_i a_i \delta_{x_i} \qquad \nu = \sum_j b_j \delta_{y_j}$$

$$\langle P, C \rangle = \sum_{ij} P_{ij} C_{ij}$$



where

$$C_{ij} = d(x_i, y_j)^p$$

subject to

$$\sum_j P_{ij} = a_i$$

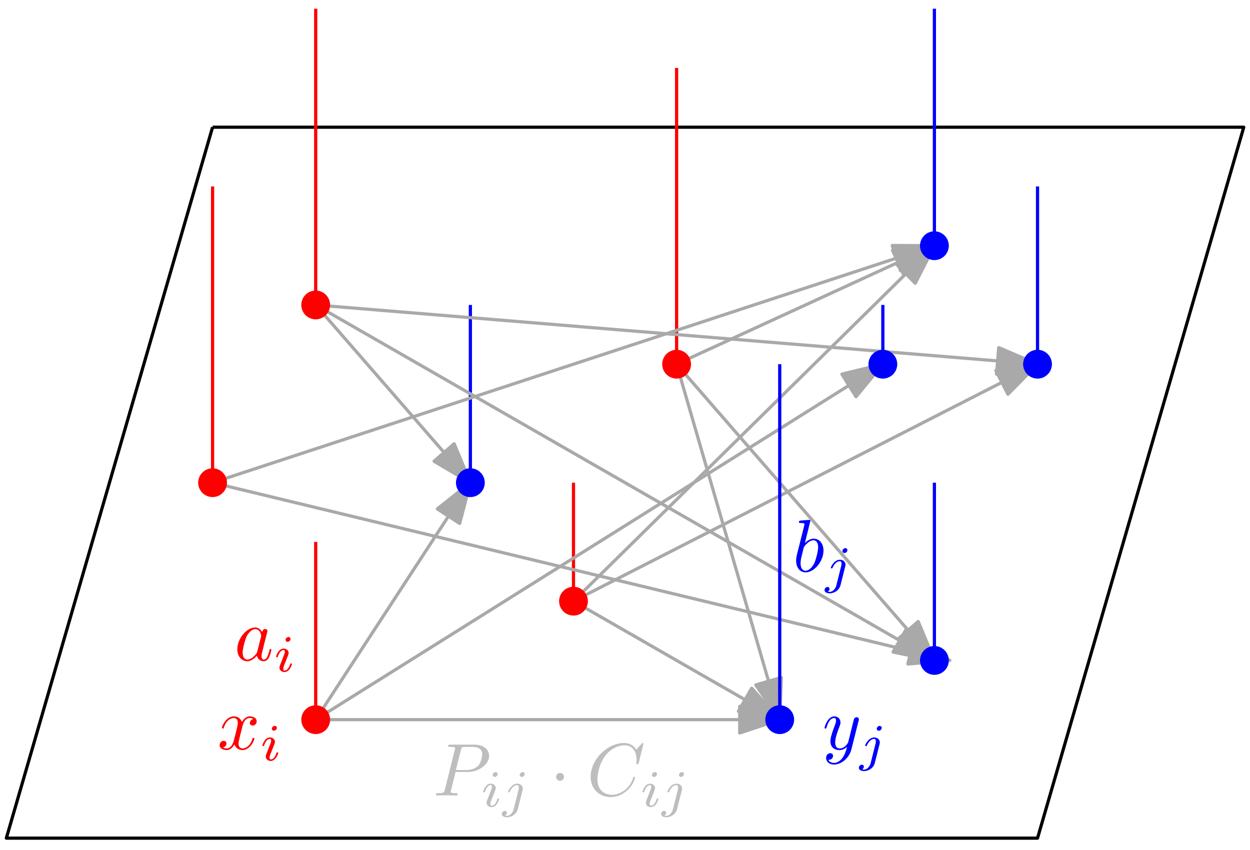
$$\sum_i P_{ij} = b_j$$

Optimal Transport - Generalities

Discrete formulation : μ and ν two probability measures

$$\mu = \sum_i a_i \delta_{x_i} \qquad \nu = \sum_j b_j \delta_{y_j}$$

$$W_p(\mu, \nu)^p = \inf_P \langle P, C \rangle = \inf_P \sum_{ij} P_{ij} C_{ij}$$



where
$C_{ij} = d(x_i, y_j)^p$

subject to

$$\sum_j P_{ij} = a_i$$

$$\sum_i P_{ij} = b_j$$

Optimal Transport - Generalities

General formulation

Consider μ, ν two probability measures on a Polish metric space $(\mathcal{X}, d)$

$\pi \in \mathcal{M}(\mathcal{X} \times \mathcal{X})$ is a *transport plan* between μ and ν if

$$\pi(A, \mathcal{X}) = \mu(A) \text{ and } \pi(\mathcal{X}, B) = \nu(B)$$

The cost of π is $C_p(\pi) := \iint_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \mathrm{d}\pi(x,y)$

and the Wasserstein-p distance between μ and ν is

$$W_p(\mu, \nu) = \left(\inf_\pi C_p(\pi)\right)^{\frac{1}{p}}$$

Optimal Transport - Generalities

Properties:

- W_p is a distance over $\{\mu \in \mathcal{P}(\mathcal{X}) : \overbrace{\int_{\mathcal{X}} d(x, x_0)^p \mathrm{d}\mu(x)}^{W_p(\mu, \delta_{x_0})^p} < \infty\}$
- It metricizes the weak convergence and the p-th moment convergence.

Optimal Transport - Generalities

Properties:

- W_p is a distance over $\{\mu \in \mathcal{P}(\mathcal{X}) : \overbrace{\int_{\mathcal{X}} d(x, x_0)^p \mathrm{d}\mu(x)}^{W_p(\mu,\delta_{x_0})^p} < \infty\}$

- It metricizes the weak convergence and the p-th moment convergence.

$$W_p(\mu_n, \mu) \to 0 \Leftrightarrow \begin{cases} \mu_n \to \mu \text{ weakly} \\ W_p(\mu_n, \delta_{x_0}) \to W_p(\mu, \delta_{x_0}) \end{cases}$$

Reminder:

- $\mu_n \to \mu$ weakly means:

for all f continuous, bounded, $\int_{\mathcal{X}} f(x)\mathrm{d}\mu_n(x) = \mu_n(f) \to \mu(f)$

- $\mu_n \to \mu$ vaguely means:

for all f continuous, compactly supported, $\mu_n(f) \to \mu(f)$

Optimal Transport - Generalities

Properties:

- W_p is a distance over $\{\mu \in \mathcal{P}(\mathcal{X}) : \overbrace{\int_{\mathcal{X}} d(x, x_0)^p \mathrm{d}\mu(x)}^{W_p(\mu,\delta_{x_0})^p} < \infty\}$

- It metricizes the weak convergence and the p-th moment convergence.

$$W_p(\mu_n, \mu) \to 0 \Leftrightarrow \begin{cases} \mu_n \to \mu \text{ weakly} \\ W_p(\mu_n, \delta_{x_0}) \to W_p(\mu, \delta_{x_0}) \end{cases}$$

- And many other nice properties:

  - Know about barycenters (Fréchet means).

  - Know the geodesics.

  - Many numerical tools (algorithms, libraries)...

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$	Optimal Transport $(\mathcal{W}^p, W_p)$

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support $(+$integer mass)

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support ($+$integer mass)
- Measures with different (potentially infinite) masses

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support
- Measures same masses

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support ($+$integer mass)
- Measures with different (potentially infinite) masses
- *Partial* matching distances (can match to the diagonal)

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support
- Measures same masses
- Exact *transportation* distances

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support ($+$integer mass)
- Measures with different (potentially infinite) masses
- *Partial* matching distances (can match to the diagonal)

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support
- Measures same masses
- Exact *transportation* distances
- Well-studied theoretically
- Efficients algorithms/libraries

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support (+integer mass)
- Measures with different (potentially infinite) masses
- *Partial* matching distances (can match to the diagonal)

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support
- Measures same masses
- Exact *transportation* distances
- Well-studied theoretically
- Efficients algorithms/libraries

Outline:

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support ($+$integer mass)
- Measures with different (potentially infinite) masses
- *Partial* matching distances (can match to the diagonal)

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support
- Measures same masses
- Exact *transportation* distances
- Well-studied theoretically
- Efficients algorithms/libraries

Outline:

- Turning d_p into an Optimal (partial) Transport problem [Divol, L, 2019]. (allow to consider more general measures)

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support (+integer mass)
- Measures with different (potentially infinite) masses
- *Partial* matching distances (can match to the diagonal)

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support
- Measures same masses
- Exact *transportation* distances
- Well-studied theoretically
- Efficients algorithms/libraries

Outline:

- Turning d_p into an Optimal (partial) Transport problem [Divol, L, 2019]. (allow to consider more general measures)
- Proving some new statistical results [Divol, L, 2019]: existence of Fréchet means, stability of linear vectorizations...

Bridging TDA and OT

Persistence Diagrams $(\mathcal{D}^p, d_p)$

- Discrete support (+integer mass)
- Measures with different (potentially infinite) masses
- *Partial* matching distances (can match to the diagonal)

Optimal Transport $(\mathcal{W}^p, W_p)$

- General support
- Measures same masses
- Exact *transportation* distances
- Well-studied theoretically
- Efficients algorithms/libraries

Outline:

- Turning d_p into an Optimal (partial) Transport problem [Divol, L, 2019]. (allow to consider more general measures)
- Proving some new statistical results [Divol, L, 2019]: existence of Fréchet means, stability of linear vectorizations...
- Deriving some efficient algorithms [L, Cuturi, Oudot, 2018]: distance and Fréchet means estimation, clustering, quantization.

Optimal Partial Transport

Global observation:

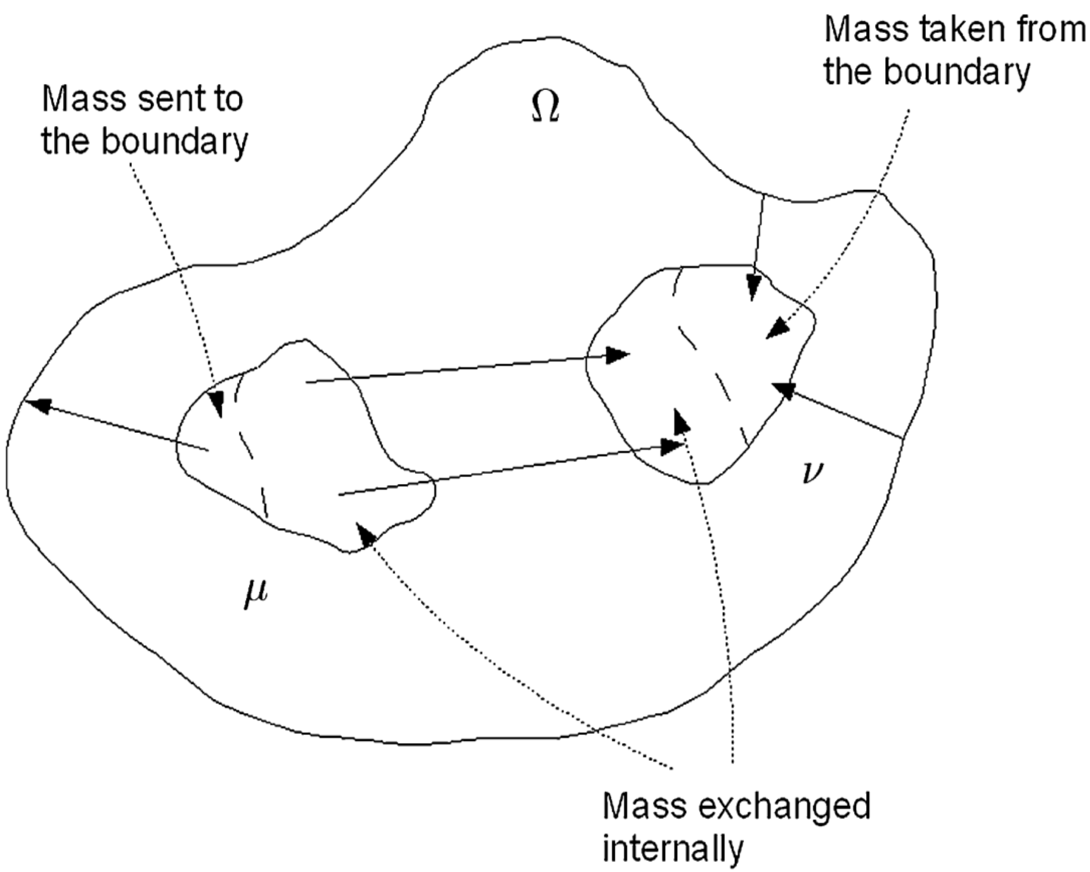
- Standard OT requires measures with the same mass
- We want to be able to handle:
    - Measures with different masses
    - Measures with infinite mass

Recent efforts were made to develop this area (eg [Chizat, 2017])

Optimal Partial Transport

A.Figalli and N.Gigli (2010)

A. Figalli, N. Gigli / J. Math. Pures Appl. 94 (2010) 107–130

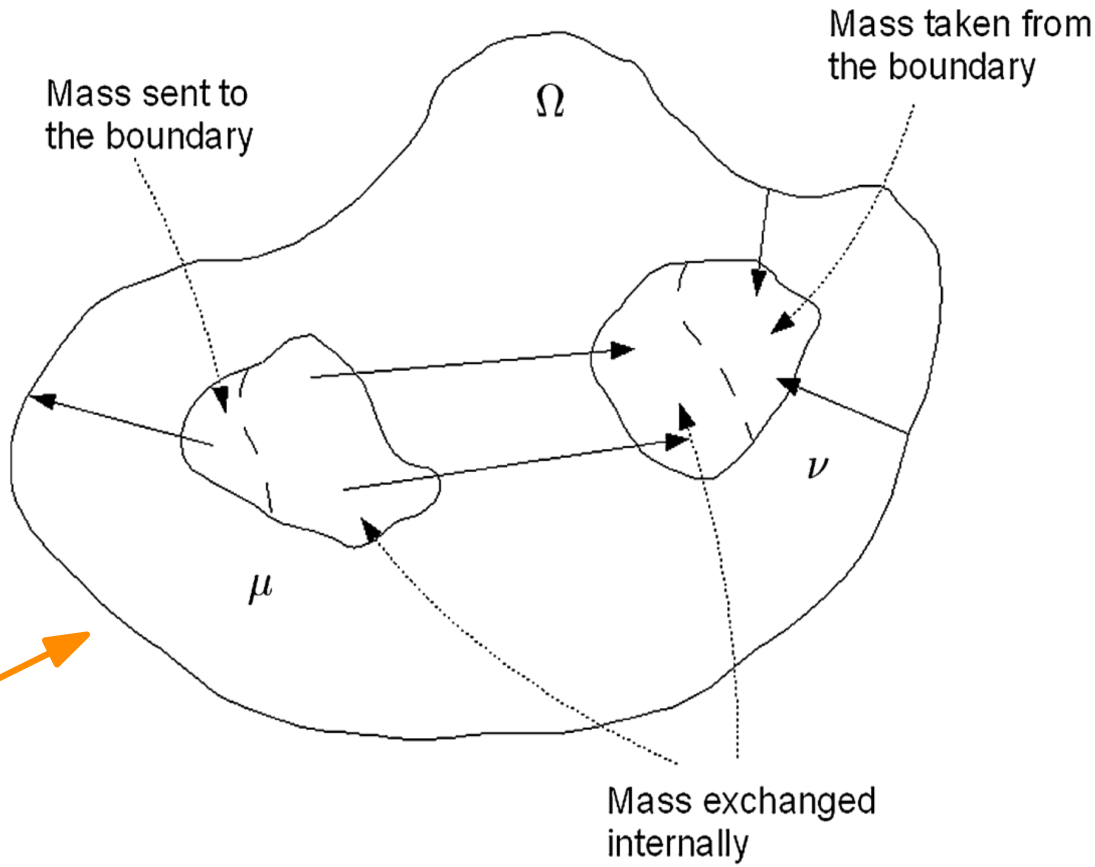


This means that we can use $\partial\Omega$ as an infinite reserve: we can 'take' as mass as we wish from the boundary, or 'give' it back some of the mass, provided we pay the transportation cost, see Fig. 1. This is why this distance is well defined for measures which do not have the same mass.

Optimal Partial Transport

A.Figalli and N.Gigli (2010)

A. Figalli, N. Gigli / J. Math. Pures Appl. 94 (2010) 107–130



bounded

This means that we can use $\partial\Omega$ as an infinite reserve: we can ‘take’ as mass as we wish from the boundary, or ‘give’ it back some of the mass, provided we pay the transportation cost, see Fig. 1. This is why this distance is well defined for measures which do not have the same mass.

Optimal Partial Transport

Core idea: Just consider sub-marginal constraints!

Given μ, ν two Radon measures on Ω, consider admissible transport plans

Let $\partial\Omega$ be the boundary of Ω, and $\overline{\Omega} = \Omega \cup \partial\Omega$

$$\pi \in \mathcal{M}(\overline{\Omega} \times \overline{\Omega}) \qquad \text{such that} \qquad \begin{aligned} \pi(A \times \overline{\Omega}) &= \mu(A) \qquad A \subset \Omega \\ \pi(\overline{\Omega} \times B) &= \nu(B) \qquad B \subset \Omega \end{aligned}$$

And then just define

$$C_p(\pi) = \iint_{\overline{\Omega}\times\overline{\Omega}} d(x,y)^p \mathrm{d}\pi(x,y)$$

$$D_p(\mu,\nu) = \left(\inf_{\pi \in \mathrm{Adm}(\mu,\nu)} C_p(\pi)\right)^{1/p}$$

Rem: measures must satisfy $\int_\Omega d(x,\partial\Omega)^p \mathrm{d}\mu(x) < +\infty$

Optimal Partial Transport

Core idea: Just consider sub-marginal constraints!

Given μ, ν two Radon measures on Ω, consider admissible transport plans

Let $\partial\Omega$ be the boundary of Ω, and $\overline{\Omega} = \Omega \cup \partial\Omega$

$$\pi \in \mathcal{M}(\overline{\Omega} \times \overline{\Omega}) \qquad \text{such that} \qquad \pi(A \times \overline{\Omega}) = \mu(A) \qquad A \subset \Omega$$

$$\pi(\overline{\Omega} \times B) = \nu(B) \qquad B \subset \Omega$$

And then just define

$$C_p(\pi) = \iint_{\overline{\Omega} \times \overline{\Omega}} d(x, y)^p \mathrm{d}\pi(x, y)$$

$$D_p(\mu, \nu) = \left(\inf_{\pi \in \mathrm{Adm}(\mu,\nu)} C_p(\pi)\right)^{1/p}$$

Rem: measures must satisfy $\int_\Omega d(x, \partial\Omega)^p \mathrm{d}\mu(x) < +\infty$

Proposition [Divol, L, 2019]:

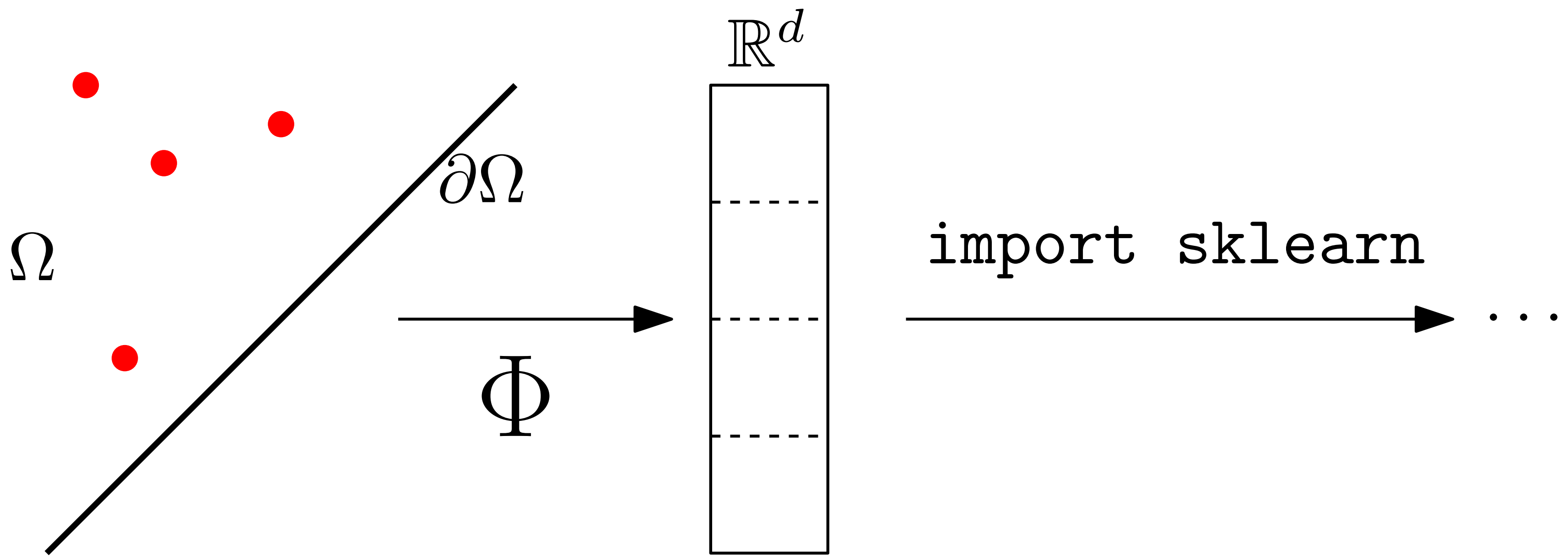
If a, b are persistence diagrams, then $D_p(a, b) = d_p(a, b)$

Stability of linear vectorizations

Stability of linear vectorizations

Context: How to use diagrams in ML pipelines?

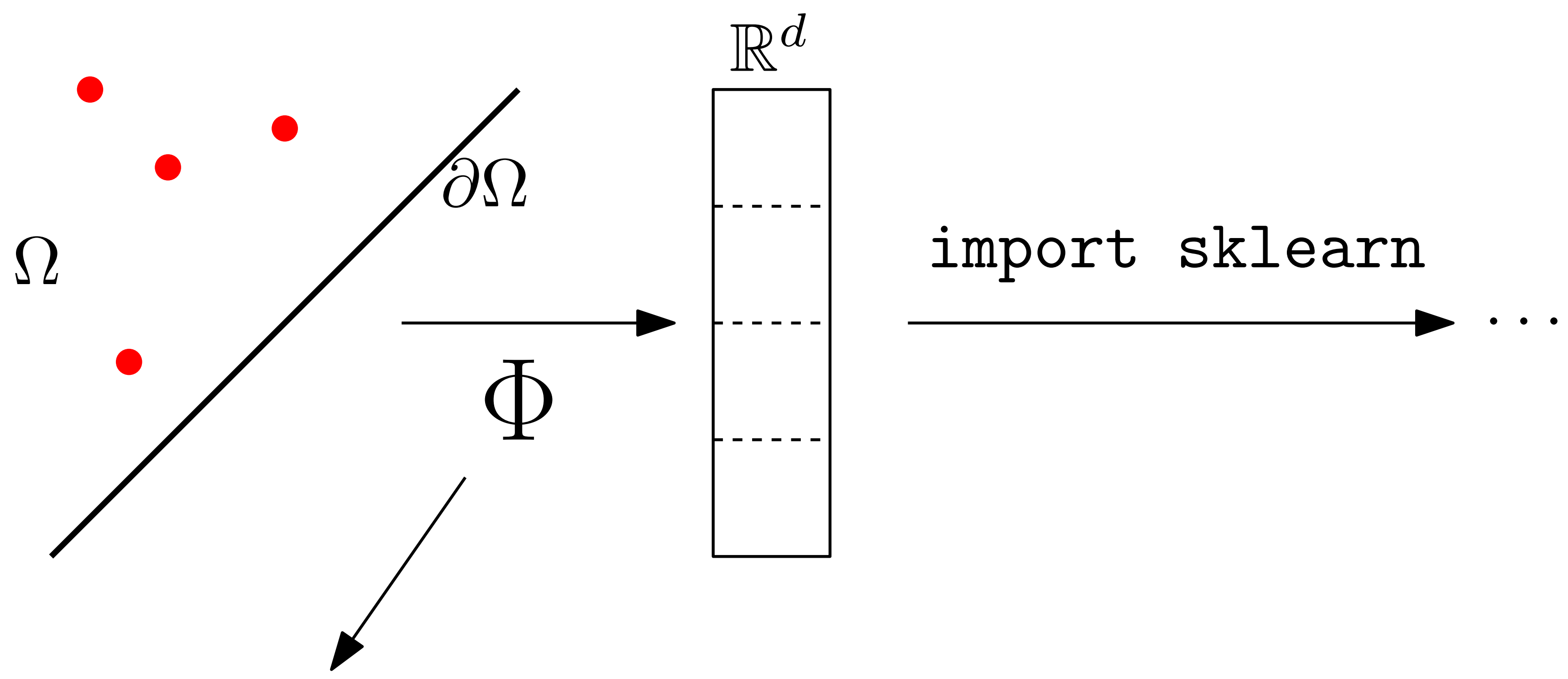
- Use kernels (not in this talk)
- Vectorize your diagrams, i.e. build $\Phi : \mathcal{D}^p \to \mathbb{R}^d$



Stability of linear vectorizations

Context: How to use diagrams in ML pipelines?

- Use kernels (not in this talk)
- Vectorize your diagrams, i.e. build $\Phi : \mathcal{D}^p \to \mathbb{R}^d$



Φ's properties? Interpretation, stability (continuity, Lipschitz)?

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

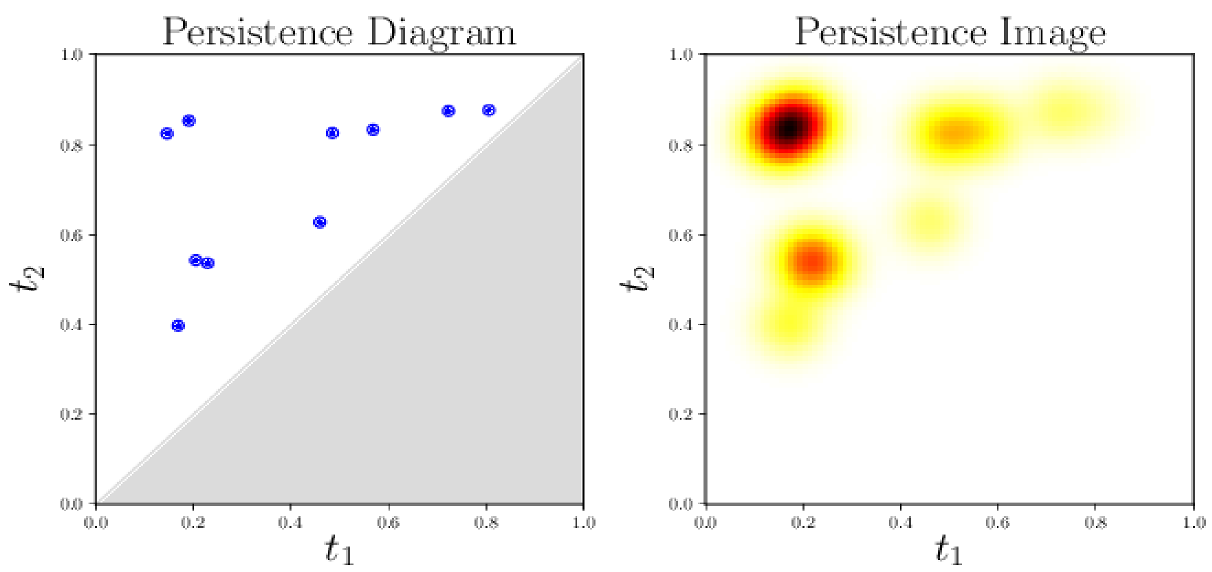
Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Example (persistence images, Adams et al, 2016):

$$f : x \mapsto d(x, \partial\Omega) \cdot \exp\left(\frac{\|x - \cdot\|^2}{2\sigma^2}\right) \text{ and } \mathcal{B} = (C_b(\mathbb{R}), \|\cdot\|_\infty)$$



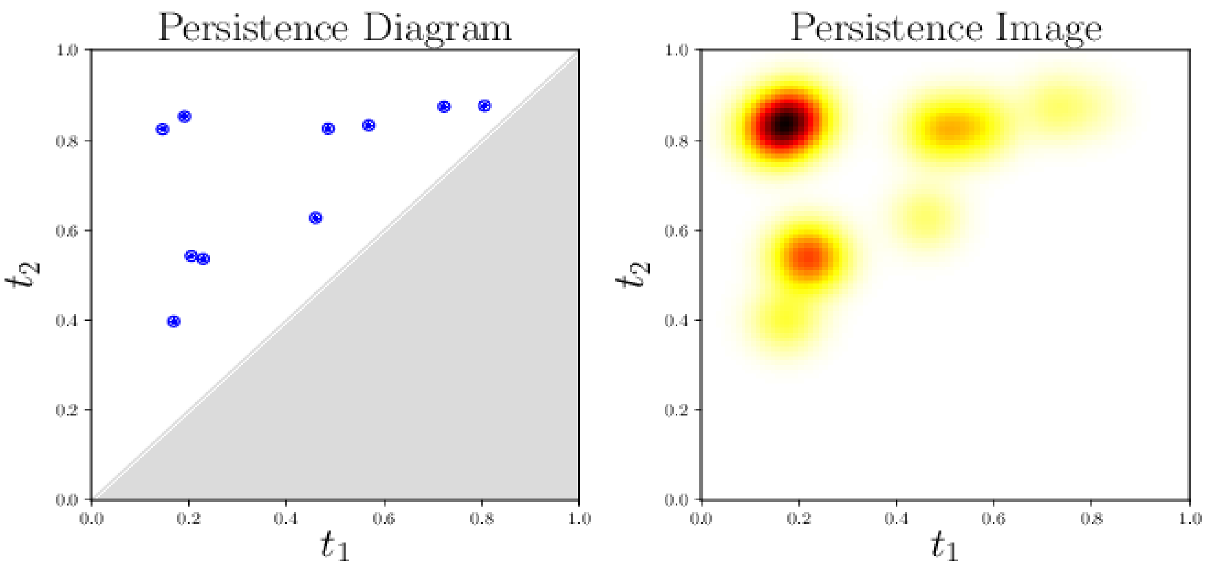
Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Example (persistence images, Adams et al, 2016):

$$f : x \mapsto d(x, \partial\Omega) \cdot \exp\left(\frac{\|x - \cdot\|^2}{2\sigma^2}\right) \text{ and } \mathcal{B} = (C_b(\mathbb{R}), \|\cdot\|_\infty)$$



Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x) d(x, \partial\Omega)^p$
where g is continuous, bounded.

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x)d(x, \partial\Omega)^p$
where g is continuous, bounded.

Idea of the proof ($\Rightarrow$): (by contradiction)

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x)d(x, \partial\Omega)^p$ where g is continuous, bounded.

Idea of the proof ($\Rightarrow$): (by contradiction)

- Consider $(x_n)_n$ such that $\frac{\|f(x_n)\|}{d(x_n,\partial\Omega)^p} \to +\infty$.
- Let $\mu_n := \frac{1}{\|f(x_n)\|}\delta_{x_n}$

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x) d(x, \partial\Omega)^p$ where g is continuous, bounded.

Idea of the proof ($\Rightarrow$): (by contradiction)

- Consider $(x_n)_n$ such that $\frac{\|f(x_n)\|}{d(x_n, \partial\Omega)^p} \to +\infty$.
- Let $\mu_n := \frac{1}{\|f(x_n)\|} \delta_{x_n}$

$D_p(\mu_n, \emptyset) = \frac{d(x_n, \partial\Omega)^p}{\|f(x_n)\|} \to 0$ but $\|\mu_n(f)\| = 1$ for all n.

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x) d(x, \partial\Omega)^p$

where g is continuous, bounded.

Idea of the proof ($\Leftarrow$):

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x) d(x, \partial\Omega)^p$ where g is continuous, bounded.

Idea of the proof ($\Leftarrow$):

- Show that $D_p(\mu_n, \mu) \to 0 \Leftrightarrow \begin{cases} \mu_n \to \mu \text{ vaguely} \\ D_p(\mu_n, \emptyset) \to d_p(\mu, \emptyset) \end{cases}$

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x) d(x, \partial\Omega)^p$ where g is continuous, bounded.

Idea of the proof ($\Leftarrow$):

- Show that $D_p(\mu_n, \mu) \to 0 \Leftrightarrow \begin{cases} \mu_n \to \mu \text{ vaguely} \\ D_p(\mu_n, \emptyset) \to d_p(\mu, \emptyset) \end{cases}$
- Deduce that $d(\cdot, \partial\Omega)^p \mu_n \to d(\cdot, \partial\Omega)^p \mu$ weakly
- Conclude that $\Phi(\mu_n) \to \Phi(\mu)$

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x)d(x, \partial\Omega)^p$
where g is continuous, bounded.

Furthermore, if $p = 1$, and if f is 1-Lipschitz continuous,

then Φ is also 1-Lipschitz continuous, ie

$$\|\Phi(\mu) - \Phi(\nu)\| = \|\mu(f) - \nu(f)\| \leq d_1(\mu, \nu)$$

Stability of linear vectorizations

Let $f : \Omega \to \mathcal{B}$ continuous, and $\Phi : \mathcal{D}^p \to \mathcal{B}$ defined by

$$\Phi(\mu) = \mu(f) = \sum_i f(x_i) \in \mathcal{B}$$

Theorem:

Φ is continuous if and only if $f(x) = g(x)d(x, \partial\Omega)^p$

where g is continuous, bounded.

Furthermore, if $p = 1$, and if f is 1-Lipschitz continuous,

then Φ is also 1-Lipschitz continuous, ie

$$\|\Phi(\mu) - \Phi(\nu)\| = \|\mu(f) - \nu(f)\| \leq d_1(\mu, \nu)$$

Idea of the proof:

Dual formulation ($\simeq$ Kantorovich-Rubinstein formula)

Barycenters in the persistence diagram space

$\mathcal{M}^p$ is a nice space (at least not so bad) for statistics:

It is complete, separable, etc.

Barycenters in the persistence diagram space

$\mathcal{M}^p$ is a nice space (at least not so bad) for statistics:

It is complete, separable, etc.

Fréchet means (aka barycenters):

Consider $b_1 \dots b_N$ a set of diagrams

Estimating their Fréchet mean consists in computing

$$\operatorname{argmin} \left\{ \mathcal{E}(a) = \frac{1}{N} \sum_{i=1}^{N} d_2(a, b_i)^2, \ a \text{ persistence diagram} \right\}.$$

Barycenters in the persistence diagram space

$\mathcal{M}^p$ is a nice space (at least not so bad) for statistics:

It is complete, separable, etc.

Fréchet means (aka barycenters):

Consider $b_1 \dots b_N$ a set of diagrams

Estimating their Fréchet mean consists in computing

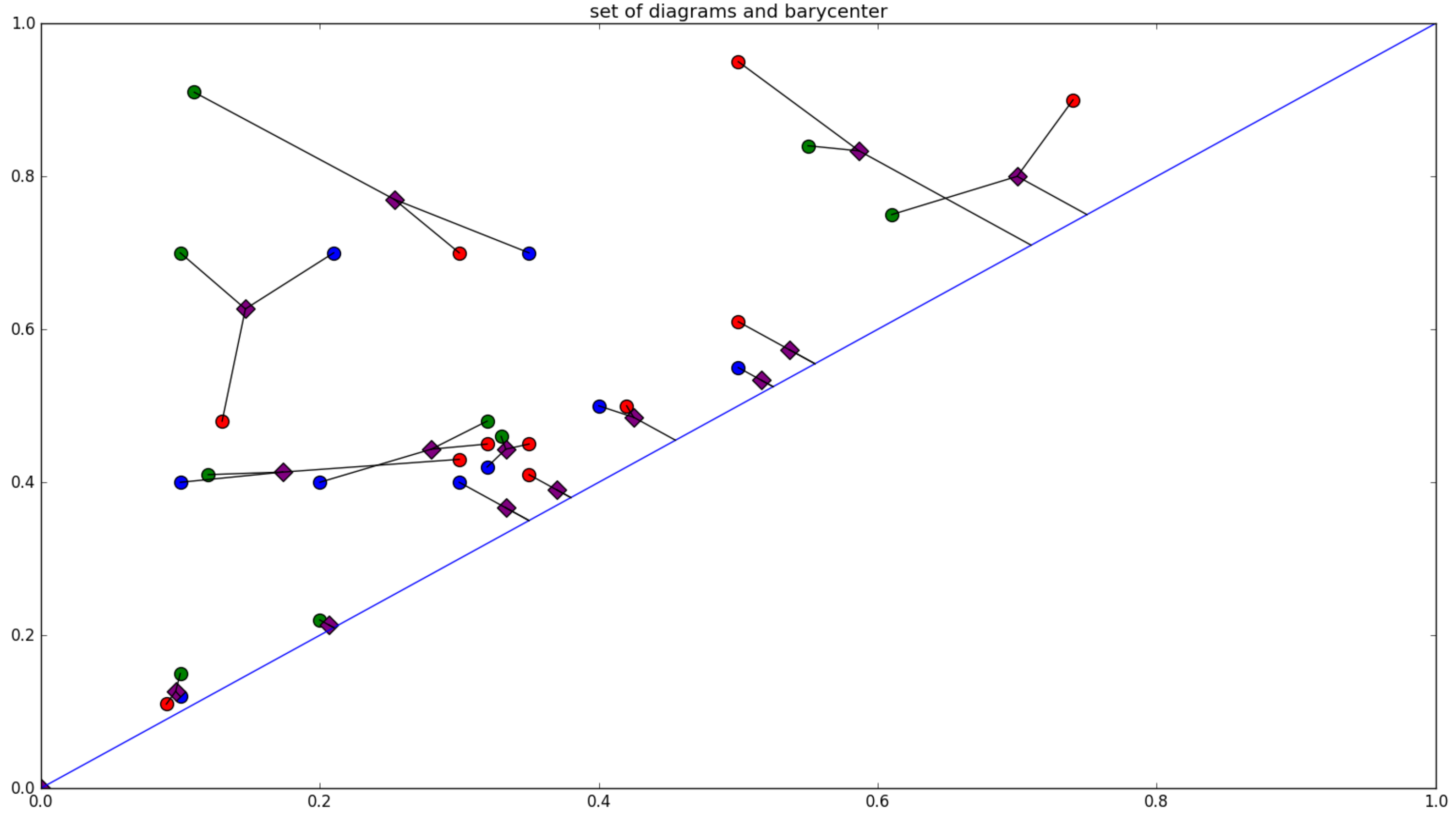
$$\operatorname{argmin} \left\{ \mathcal{E}(a) = \frac{1}{N} \sum_{i=1}^{N} d_2(a, b_i)^2, \ a \text{ persistence diagram} \right\}.$$

First results [Turner et al. 2013]:

- $\mathcal{E}$ is not convex. It admits global (and local) minimizers
- Local minimizers can be computed (expensive)

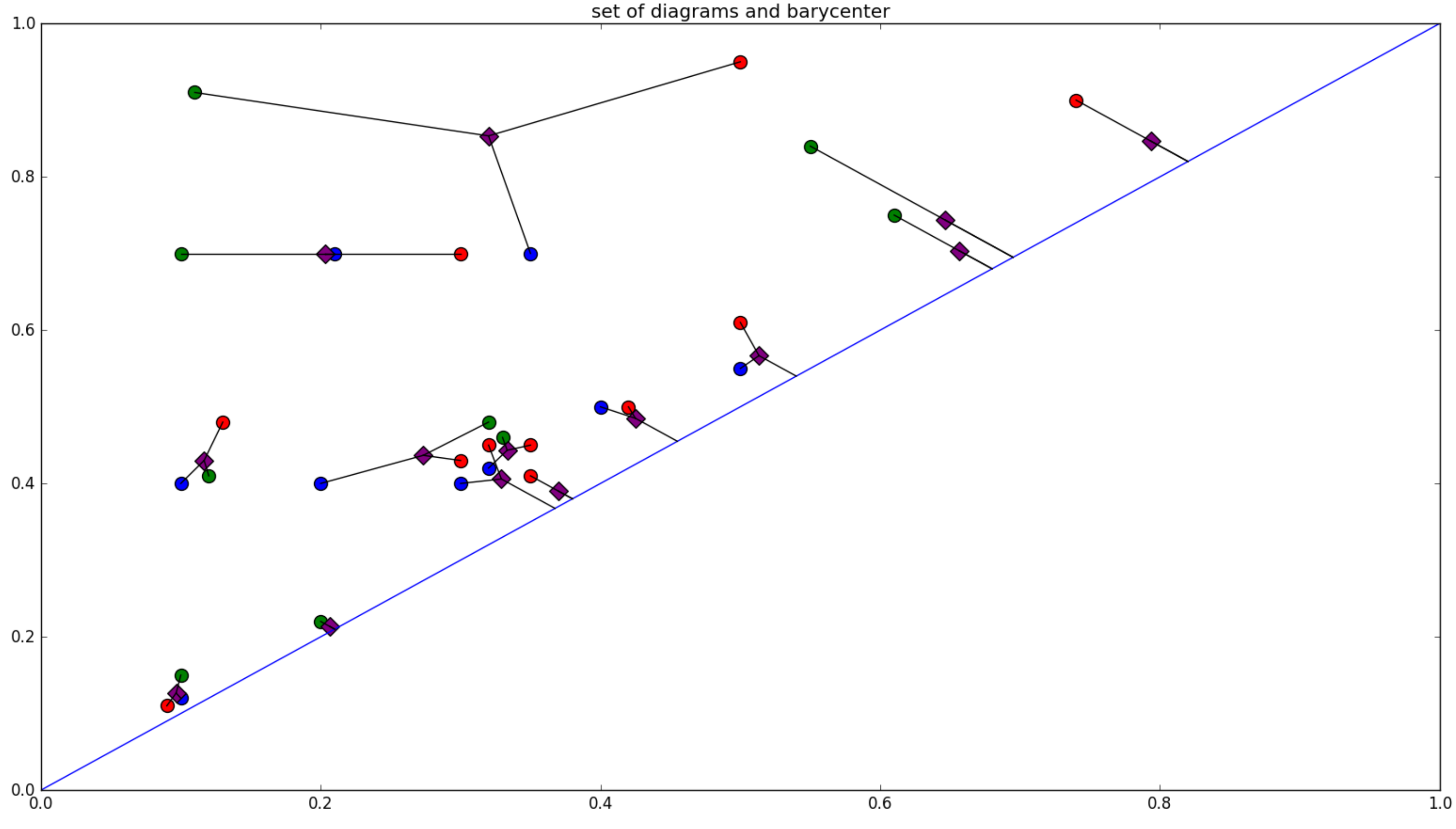
Barycenters in the persistence diagram space

Estimation of barycenters (and issues of non-convexity)



Barycenters in the persistence diagram space

Estimation of barycenters (and issues of non-convexity)



Barycenters in the persistence diagram space

$\mathcal{M}^p$ is a nice space (at least not so bad) for statistics:

It is complete, separable, etc.

Fréchet means (aka barycenters):

Consider $b_1 \dots b_N$ a set of diagrams

Estimating their Fréchet mean consists in computing

$$\operatorname{argmin}\left\{\mathcal{E}(a) = \frac{1}{N}\sum_{i=1}^{N} d_2(a, b_i)^2,\ a \ \underset{\text{any Radon measure}}{\cancel{\text{persistence diagram}}}\right\}.$$

Barycenters in the persistence diagram space

$\mathcal{M}^p$ is a nice space (at least not so bad) for statistics:

It is complete, separable, etc.

Fréchet means (aka barycenters):

Consider $b_1 \dots b_N$ a set of diagrams

Estimating their Fréchet mean consists in computing

$$\operatorname{argmin}\left\{\mathcal{E}(a) = \frac{1}{N}\sum_{i=1}^{N} d_2(a, b_i)^2,\ a \underbrace{\cancel{\text{ persistence diagram}}}_{\text{any Radon measure}}\right\}.$$

Properties [Divol, L, 2019]

- $\mathcal{E}$ is now convex, admits global minimizers.
- Some of them are actual diagrams.

Barycenters in the persistence diagram space

$\mathcal{M}^p$ is a nice space (at least not so bad) for statistics:

It is complete, separable, etc.

Fréchet means (aka barycenters):

Consider $b_1 \dots b_N$ a set of diagrams

Estimating their Fréchet mean consists in computing

$$\operatorname{argmin} \left\{ \mathcal{E}(a) = \frac{1}{N} \sum_{i=1}^{N} d_2(a, b_i)^2, \ a \text{ \cancel{persistence diagram}} \underset{\text{any Radon measure}}{} \right\}.$$

Properties [Divol, L, 2019]

- $\mathcal{E}$ is now convex, admits global minimizers.
- Some of them are actual diagrams.

Numerical considerations [L, Cuturi, Oudot, 2018]

- These can be approximated efficiently (Sinkhorn algorithm).

Conclusion

Take home messages:

- TDA (at least PDs) can be formulated as an OT pbm
- This formalism has theoretical and numerical strengths

Some other applications / links:

- Other results in the space of PDs:
  - Topological stability of random processes
  - Geodesics
- Sinkhorn divergences (Genevay et al. 2018)
- Semi-discrete transport
- Kernel for persistence diagrams