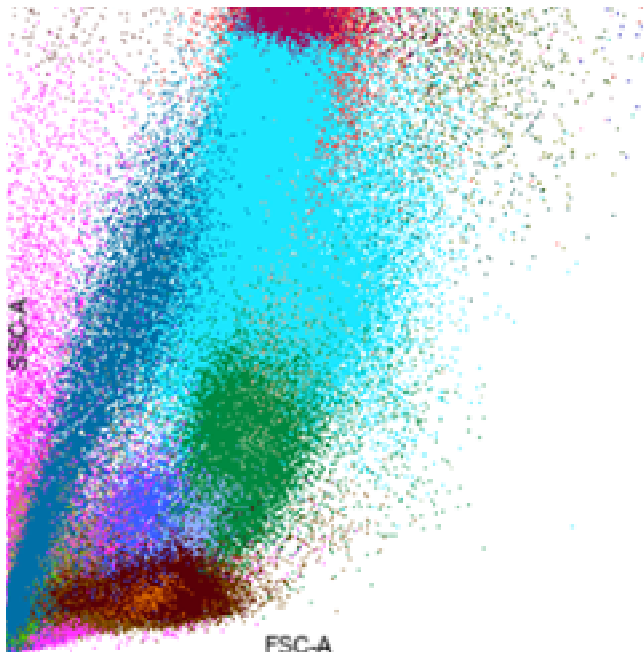
Montpellier - June, 9 2021

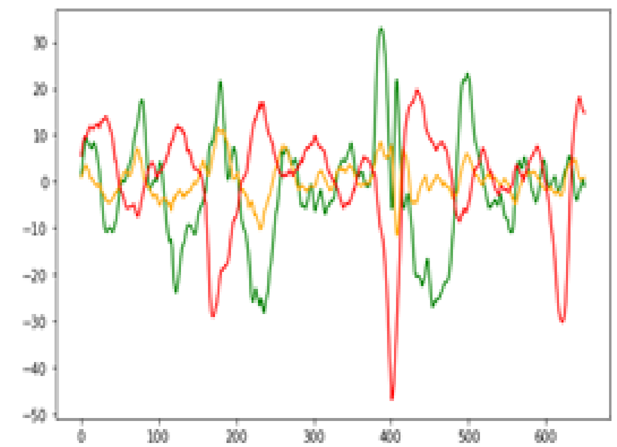
A framework to differentiate persistent homology with applications in Machine Learning and Statistics

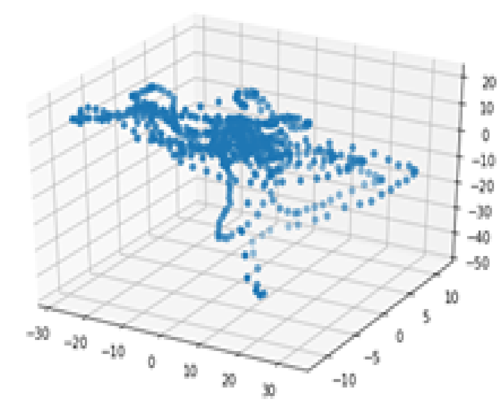
Frédéric Chazal
DataShape team
Inria & Laboratoire de Mathématiques d'Orsay

What is Topological Data Analysis (TDA) ?



[Cell population - cytometry - MetaFora courtesy]

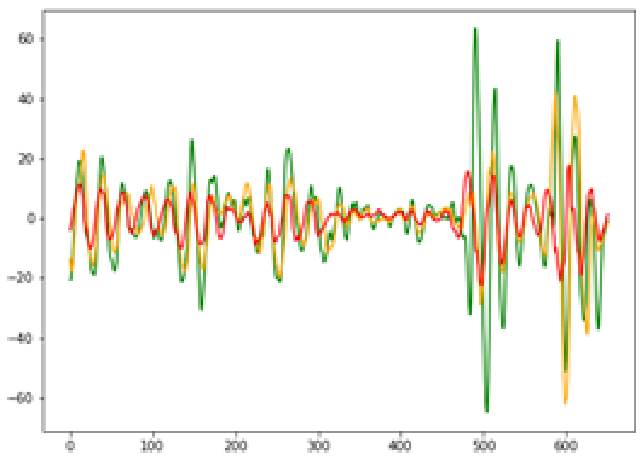


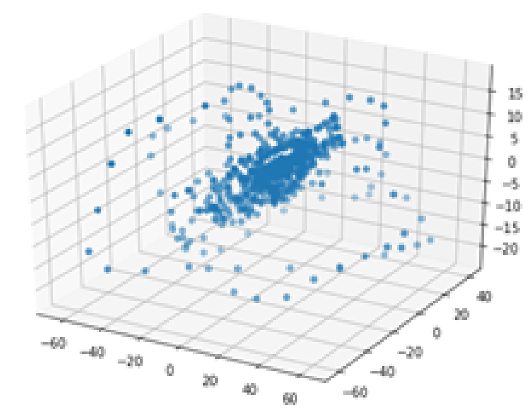


[Sensors (Sysnav courtesy)]

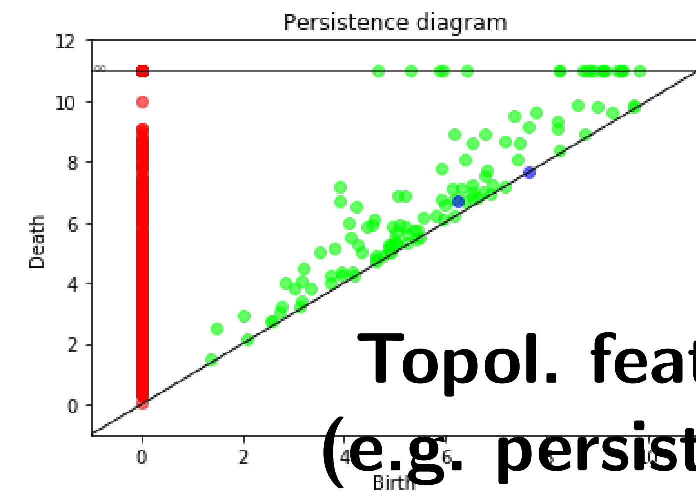
Modern data carry complex, but important, geometric/topological structure !

What is Topological Data Analysis (TDA) ?





Data

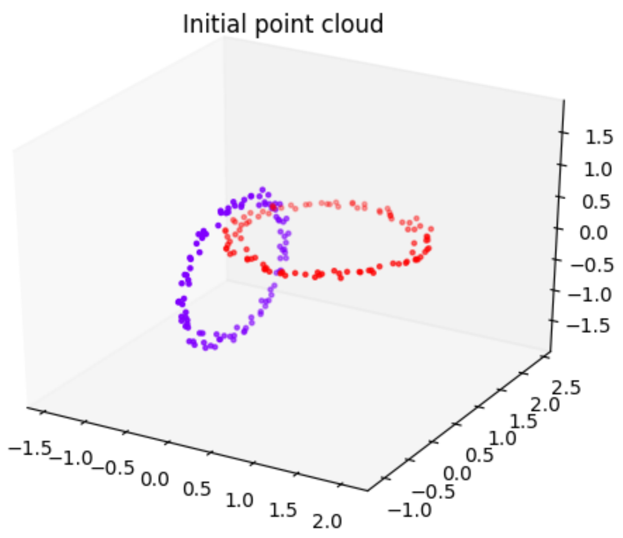


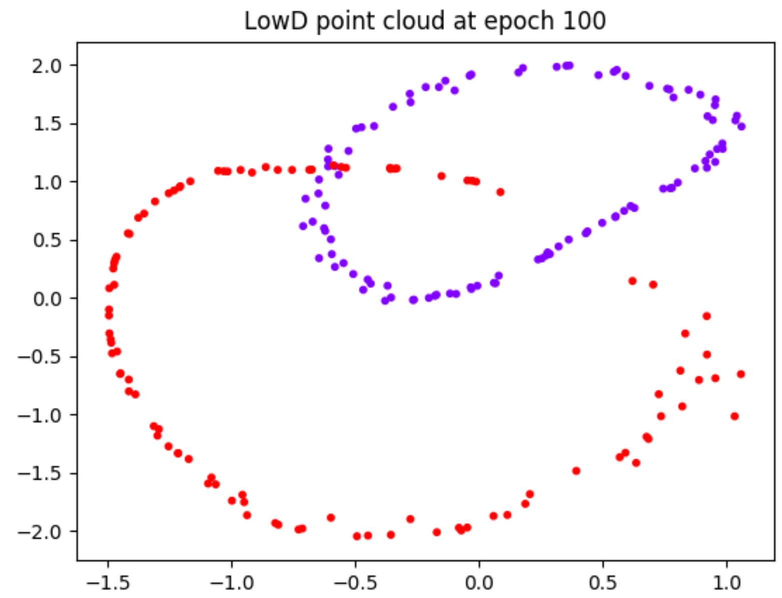
Topol. features (e.g. persistence)

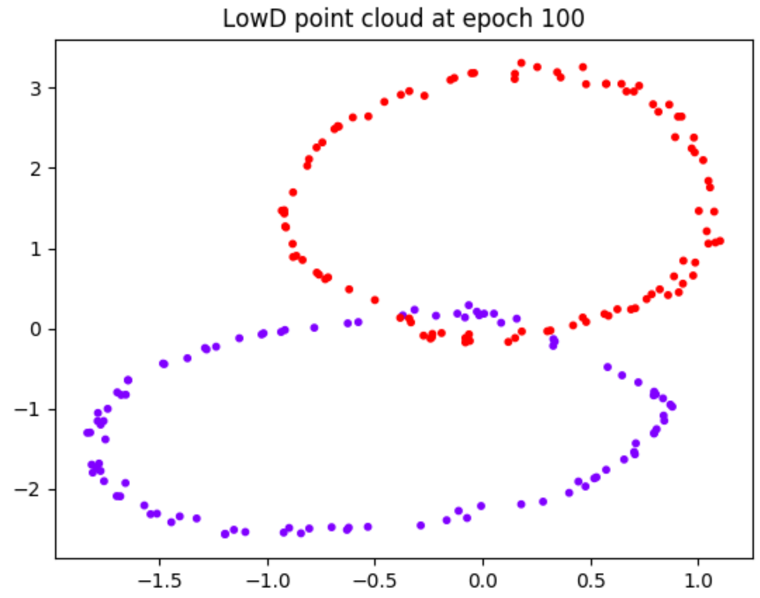
Topological Data Analysis (TDA) is a recent field whose aim is to :

- infer relevant topological and geometric features from complex data,
- take advantage of topological/geometric information for further Data Analysis, Machine Learning and AI tasks :
  - using topological features in ML pipelines,
  - taking advantage of topological information to improve ML pipelines (e.g. topological losses).

Example : dimensionality reduction







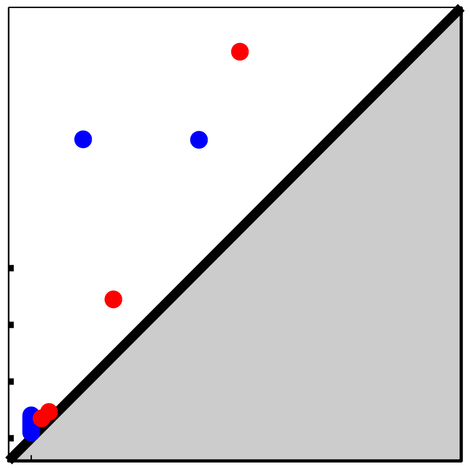
Input : 2 sampled circles in $\mathbb{R}^9$

Dim reduction in $\mathbb{R}^2$ without topol. constraint

Dim reduction in $\mathbb{R}^2$ with topol. constraint

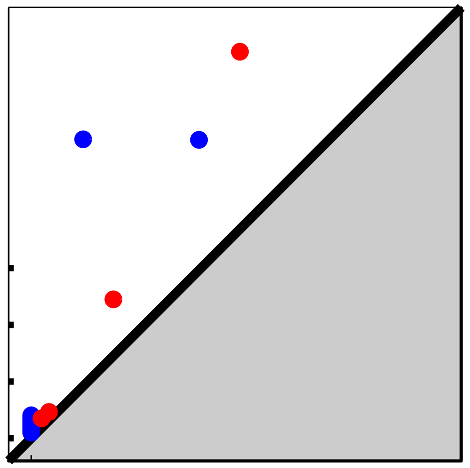
Two related general questions

$$\arg\min f(\quad)?$$



1. How to minimize functions depending of persistence diagrams (e.g. total persistence) ?
2. Can we understand the average behavior of random persistence diagrams ?

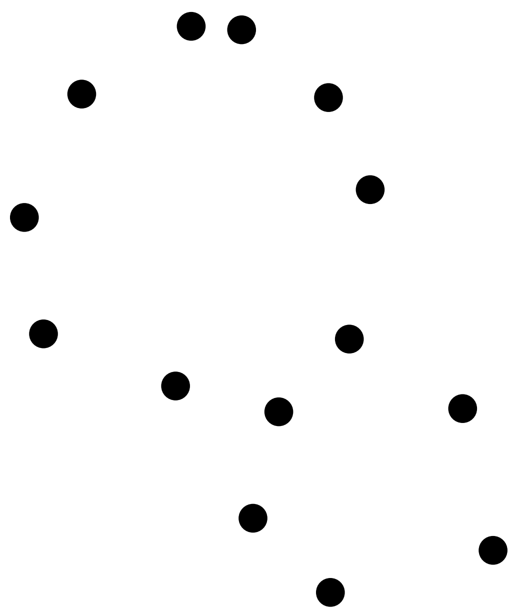
Two related general questions

$\arg\min f($  $)$?

1. How to minimize functions depending of persistence diagrams (e.g. total persistence) ?
2. Can we understand the average behavior of random persistence diagrams ?

$\rightarrow$ Both need to understand the "differentiability of persistence"

Simplicial complexes and filtrations

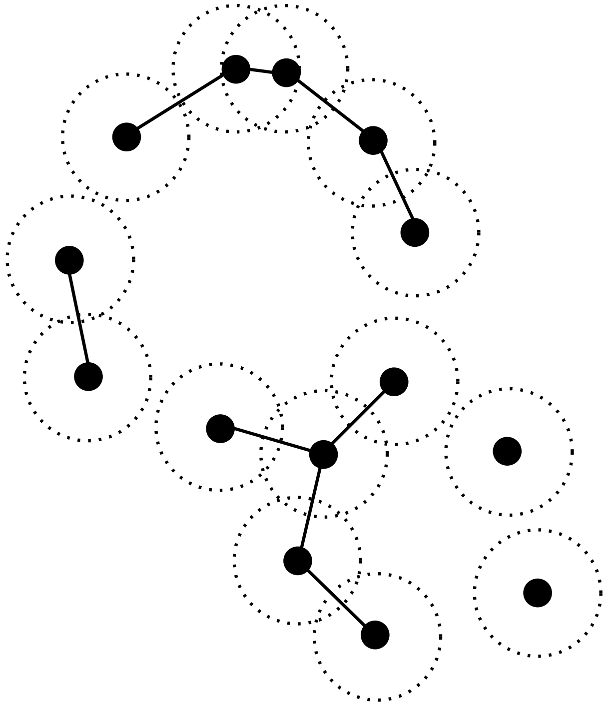


Given a set V, a simplicial complex K is a collection of finite subsets of V s. t.

- $\{v\} \in K$ for any $v \in V$,
- if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$.

Given K and $R \subseteq \mathbb{R}$, a filtration of K is an increasing sequence $(K_r)_{r\in R}$ of subcomplexes of K with respect to the inclusion such that $\bigcup_{r\in R} K_r = K$.

Simplicial complexes and filtrations

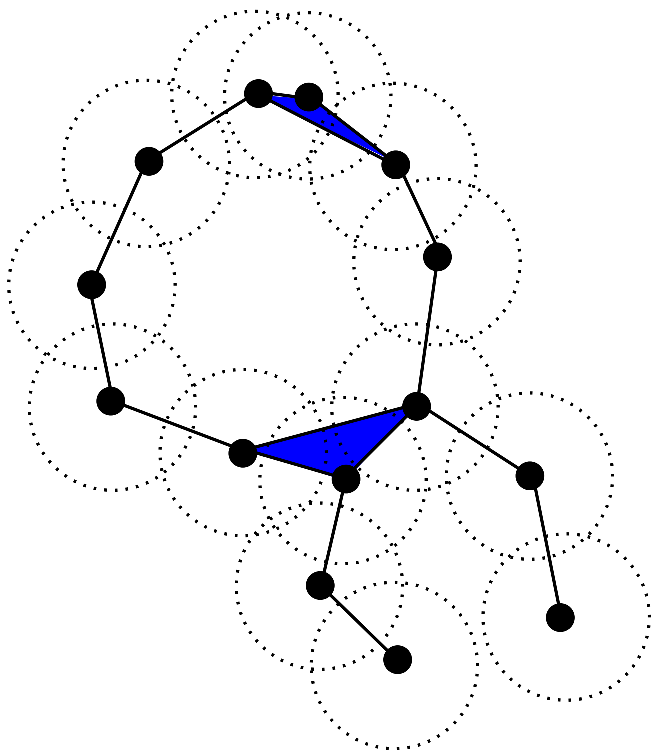


Given a set V, a simplicial complex K is a collection of finite subsets of V s. t.

- $\{v\} \in K$ for any $v \in V$,
- if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$.

Given K and $R \subseteq \mathbb{R}$, a filtration of K is an increasing sequence $(K_r)_{r \in R}$ of subcomplexes of K with respect to the inclusion such that $\bigcup_{r \in R} K_r = K$.

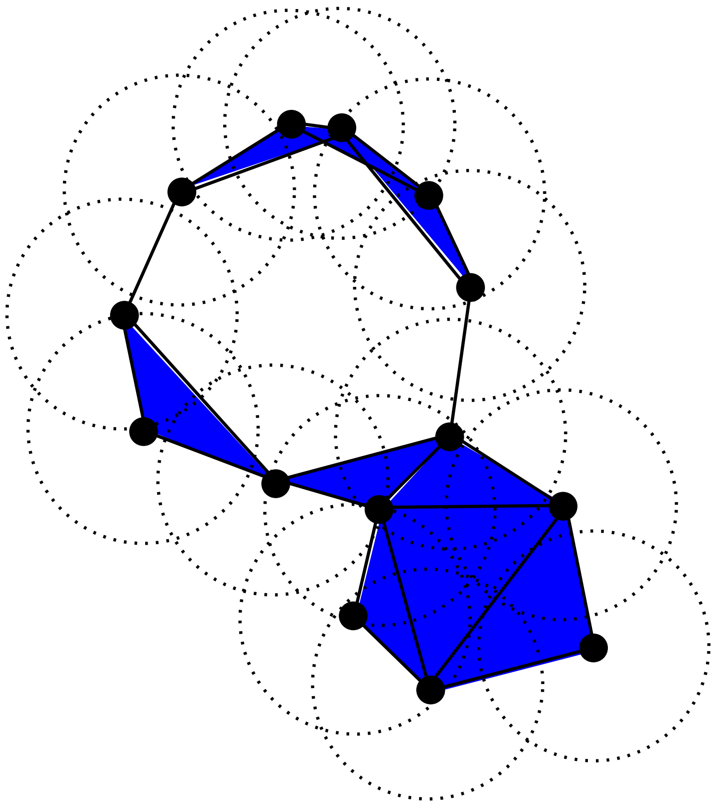
Simplicial complexes and filtrations



Given a set V, a simplicial complex K is a collection of finite subsets of V s. t.
- $\{v\} \in K$ for any $v \in V$,
- if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$.

Given K and $R \subseteq \mathbb{R}$, a filtration of K is an increasing sequence $(K_r)_{r \in R}$ of subcomplexes of K with respect to the inclusion such that $\bigcup_{r \in R} K_r = K$.

Simplicial complexes and filtrations

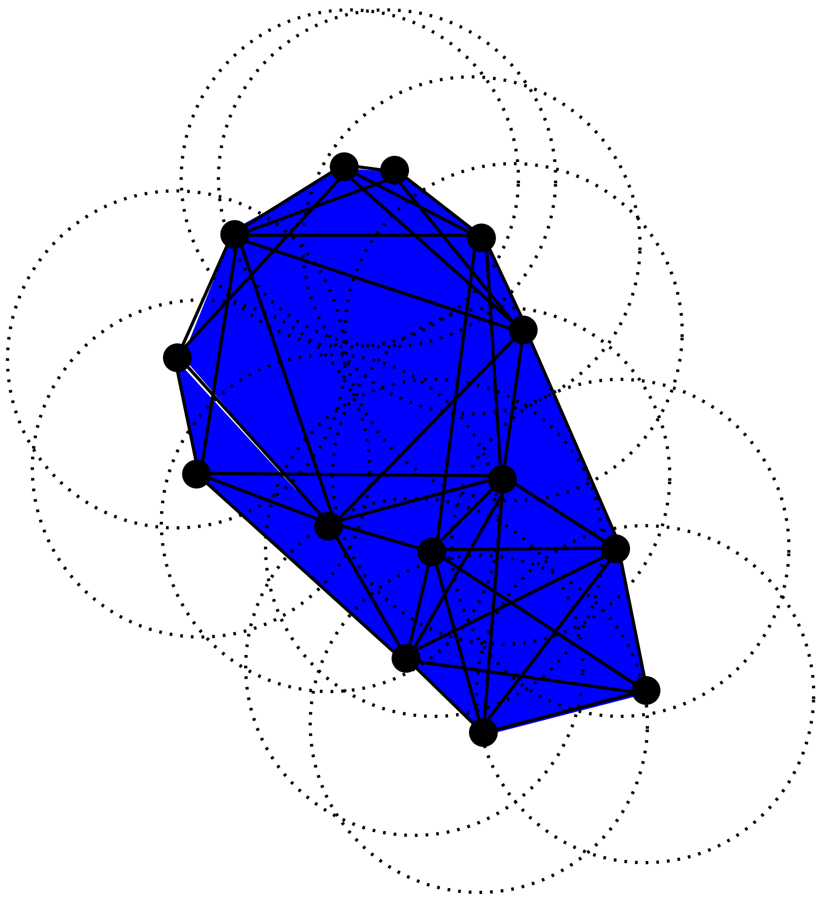


Given a set V, a simplicial complex K is a collection of finite subsets of V s. t.

- $\{v\} \in K$ for any $v \in V$,
- if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$.

Given K and $R \subseteq \mathbb{R}$, a filtration of K is an increasing sequence $(K_r)_{r \in R}$ of subcomplexes of K with respect to the inclusion such that $\bigcup_{r \in R} K_r = K$.

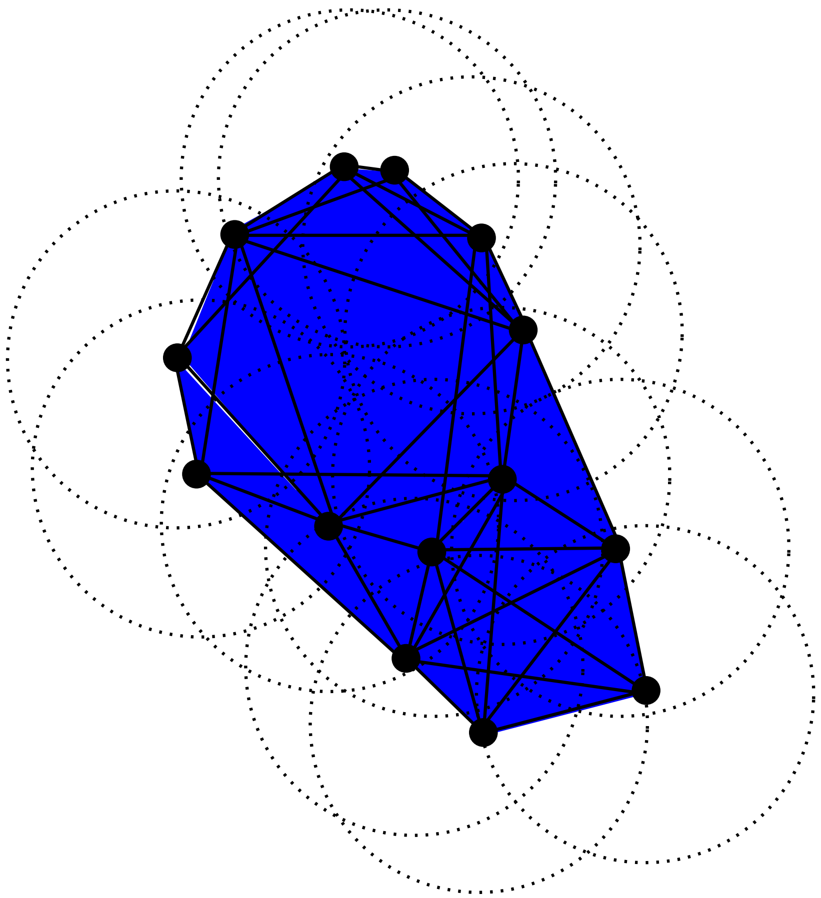
Simplicial complexes and filtrations



Given a set V, a simplicial complex K is a collection of finite subsets of V s. t.
- $\{v\} \in K$ for any $v \in V$,
- if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$.

Given K and $R \subseteq \mathbb{R}$, a filtration of K is an increasing sequence $(K_r)_{r \in R}$ of subcomplexes of K with respect to the inclusion such that $\bigcup_{r \in R} K_r = K$.

Simplicial complexes and filtrations



Given a set V, a simplicial complex K is a collection of finite subsets of V s. t.

- $\{v\} \in K$ for any $v \in V$,
- if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$.

Given K and $R \subseteq \mathbb{R}$, a filtration of K is an increasing sequence $(K_r)_{r \in R}$ of subcomplexes of K with respect to the inclusion such that $\bigcup_{r \in R} K_r = K$.

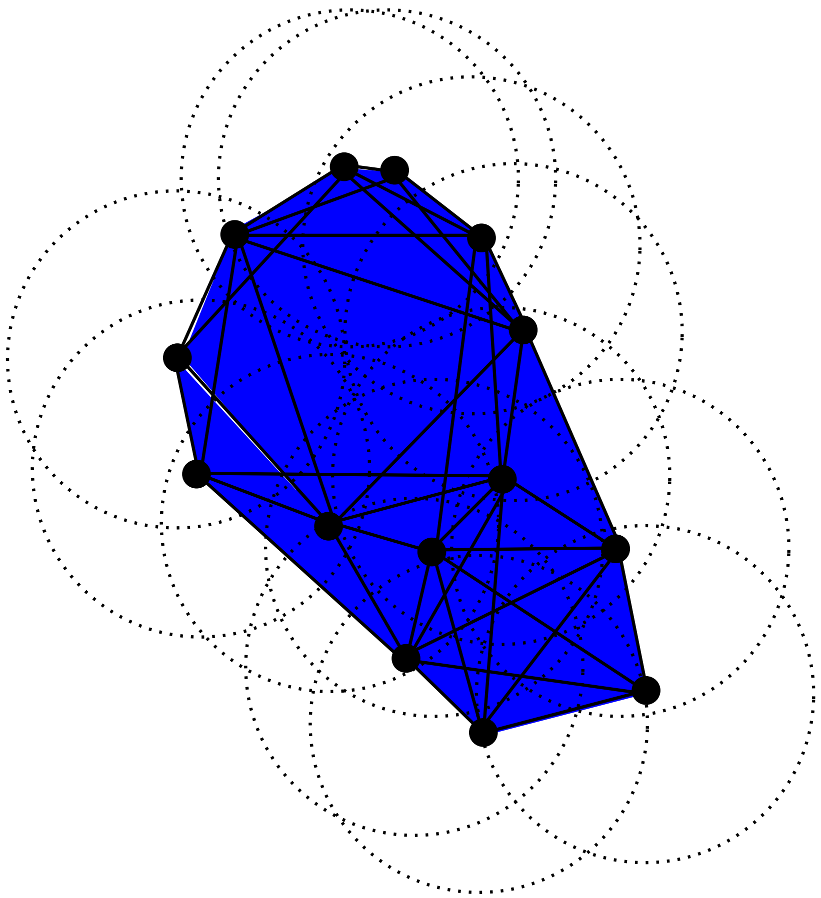
To $\sigma \in K$, one can associate $\Phi_\sigma = \inf\{r \in R : \sigma \in K_r\}$

$\Rightarrow$ A filtration of K is a $|K|$-dimensional vector

$$\Phi = (\Phi_\sigma)_{\sigma \in K} \in \mathbb{R}^{|K|} \text{ s. t. } \tau \subseteq \sigma \Rightarrow \Phi_\tau \leq \Phi_\sigma$$

The set $\mathrm{Filt}_K \subset \mathbb{R}^{|K|}$ of the vectors in $\mathbb{R}^{|K|}$ defining a filtration on K is semi-algebraic.

Simplicial complexes and filtrations



Given a set V, a simplicial complex K is a collection of finite subsets of V s. t.

- $\{v\} \in K$ for any $v \in V$,
- if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$.

Given K and $R \subseteq \mathbb{R}$, a filtration of K is an increasing sequence $(K_r)_{r \in R}$ of subcomplexes of K with respect to the inclusion such that $\bigcup_{r \in R} K_r = K$.

Definition : Let K be a simplicial complex and A a set. A map $\Phi\colon A \to \mathbb{R}^{|K|}$ is said to be a parametrized family of filtrations if for any $x \in A$ and $\sigma, \tau \in K$ with $\tau \subseteq \sigma$, one has $\Phi_\tau(x) \leq \Phi_\sigma(x)$.

Persistent homology computation

Let K be a finite filtered simplicial complex and let $\sigma_1 \preceq \cdots \preceq \sigma_{|K|}$ the simplices of K ordered according the increasing entries of $\Phi = (\Phi_\sigma)_{\sigma \in K} \in \mathbb{R}^{|K|}$

Persistent homology computation

Let K be a finite filtered simplicial complex and let $\sigma_1 \preceq \cdots \preceq \sigma_{|K|}$ the simplices of K ordered according the increasing entries of $\Phi = (\Phi_\sigma)_{\sigma \in K} \in \mathbb{R}^{|K|}$

Process the simplices according to their order of entrance in the filtration :

Let $k = \dim \sigma_i$ and denote $K_{i-1} = \cup_{l=1}^{i-1} \sigma_l$

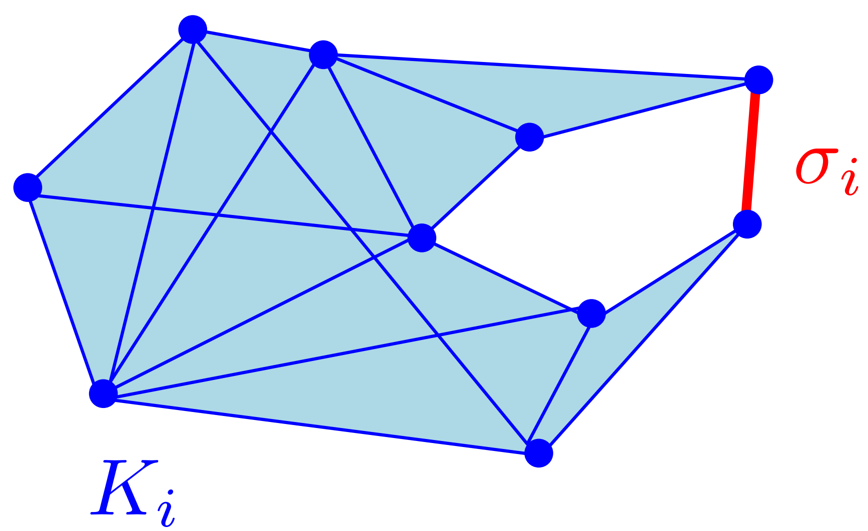
Persistent homology computation

Let K be a finite filtered simplicial complex and let $\sigma_1 \preceq \cdots \preceq \sigma_{|K|}$ the simplices of K ordered according the increasing entries of $\Phi = (\Phi_\sigma)_{\sigma \in K} \in \mathbb{R}^{|K|}$

Process the simplices according to their order of entrance in the filtration :

Let $k = \dim \sigma_i$ and denote $K_{i-1} = \cup_{l=1}^{i-1} \sigma_l$

Case 1 : adding σ_i to K_{i-1} creates a new k-dimensional topological feature in K_i (new homology class in H_k).



$\Rightarrow$ the birth of a k-dim feature is registered.

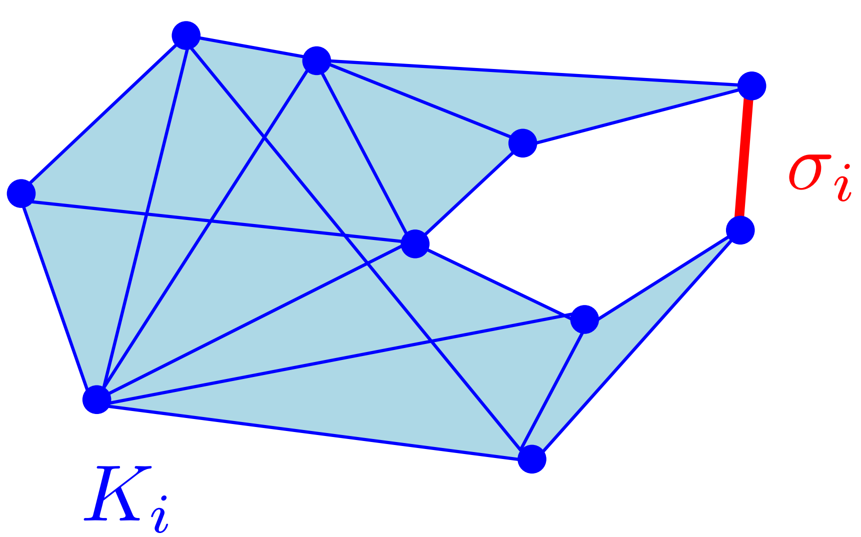
Persistent homology computation

Let K be a finite filtered simplicial complex and let $\sigma_1 \preceq \cdots \preceq \sigma_{|K|}$ the simplices of K ordered according the increasing entries of $\Phi = (\Phi_\sigma)_{\sigma \in K} \in \mathbb{R}^{|K|}$

Process the simplices according to their order of entrance in the filtration :

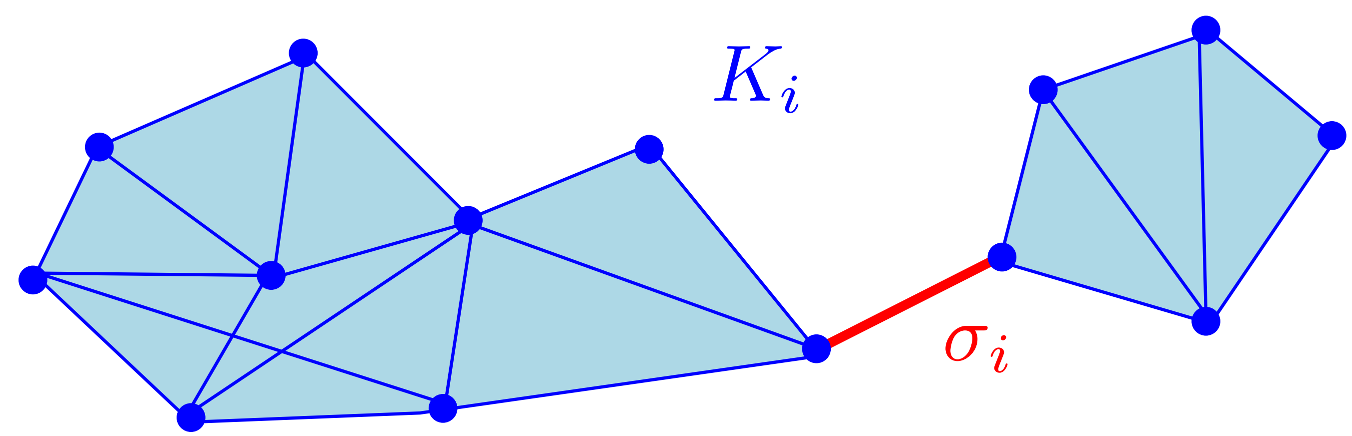
Let $k = \dim \sigma_i$ and denote $K_{i-1} = \cup_{l=1}^{i-1} \sigma_l$

Case 1 : adding σ_i to K_{i-1} creates a new k-dimensional topological feature in K_i (new homology class in H_k).



$\Rightarrow$ the birth of a k-dim feature is registered.

Case 2 : adding σ_i to K_{i-1} kills a $(k-1)$-dimensional topological feature in K_i (homology class in H_{k-1}).



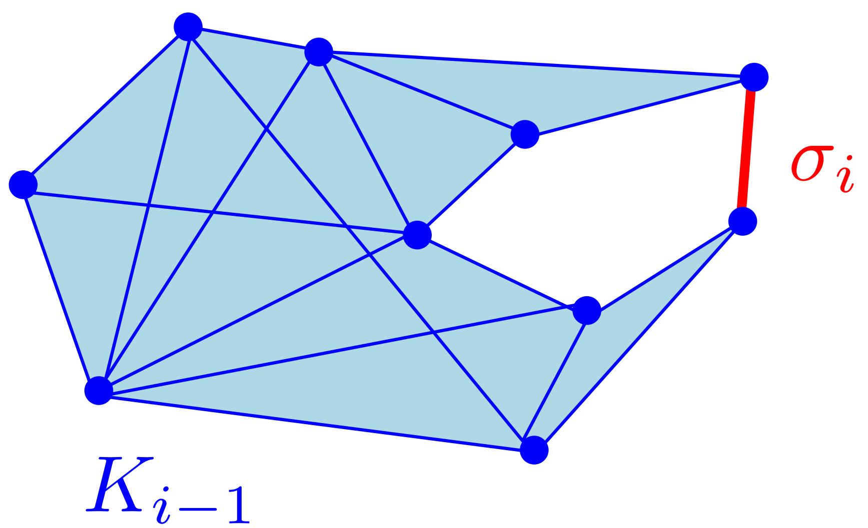
$\Rightarrow$ persistence algo. pairs the simplex σ_i to the simplex $\sigma_{l(i)}$ that gave birth to the killed feature.

Persistent homology computation

Process the simplices according to their order of entrance in the filtration :

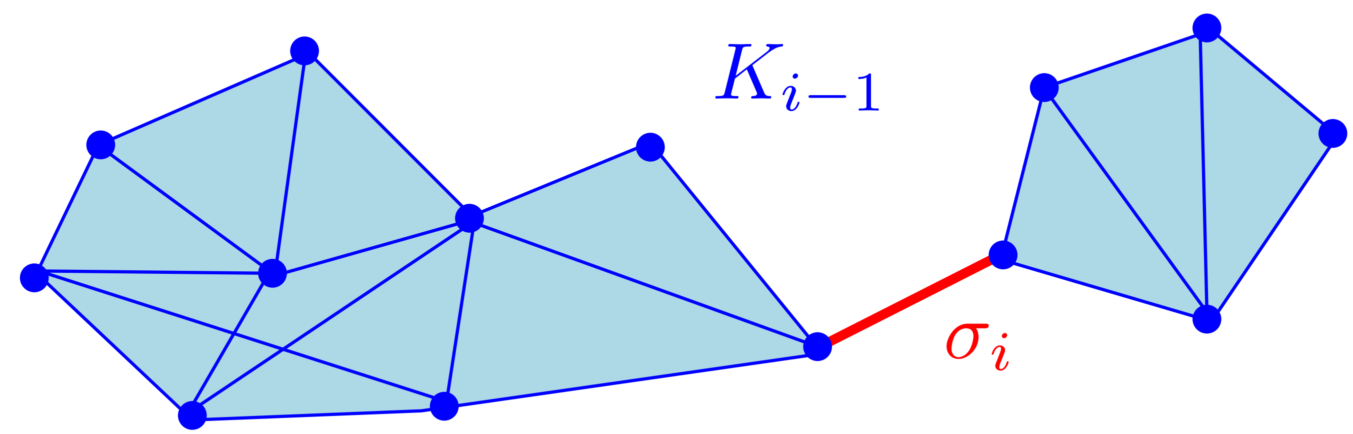
Let $k = \dim \sigma_i$ and denote $K_{i-1} = \cup_{l=1}^{i-1} \sigma_l$

Case 1 : adding σ_i to K_{i-1} creates a new k-dimensional topological feature in Ki (new homology class in H_k).



⇒ the birth of a k-dim feature is registered.

Case 2 : adding σ_i to K_{i-1} kills a $(k-1)$-dimensional topological feature in K_i (homology class in H_{k-1}).



⇒ persistence algo. pairs the simplex σ_i to the simplex $\sigma_{l(i)}$ that gave birth to the killed feature.

→ $(\sigma_{l(i)}, \sigma_i)$: persistence pair

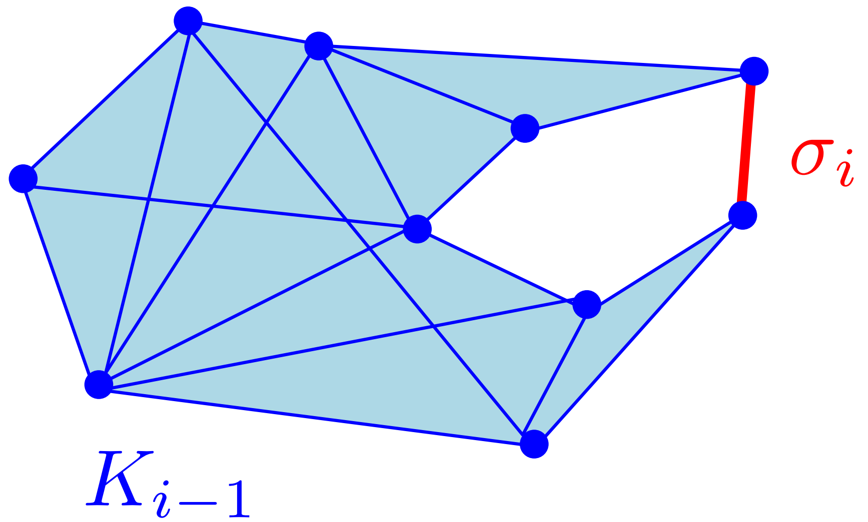
→ $(\Phi_{\sigma_{l(i)}}, \Phi_{\sigma_i}) \in \mathbb{R}^2$: point in the persistence diagram

Persistent homology computation

Process the simplices according to their order of entrance in the filtration :

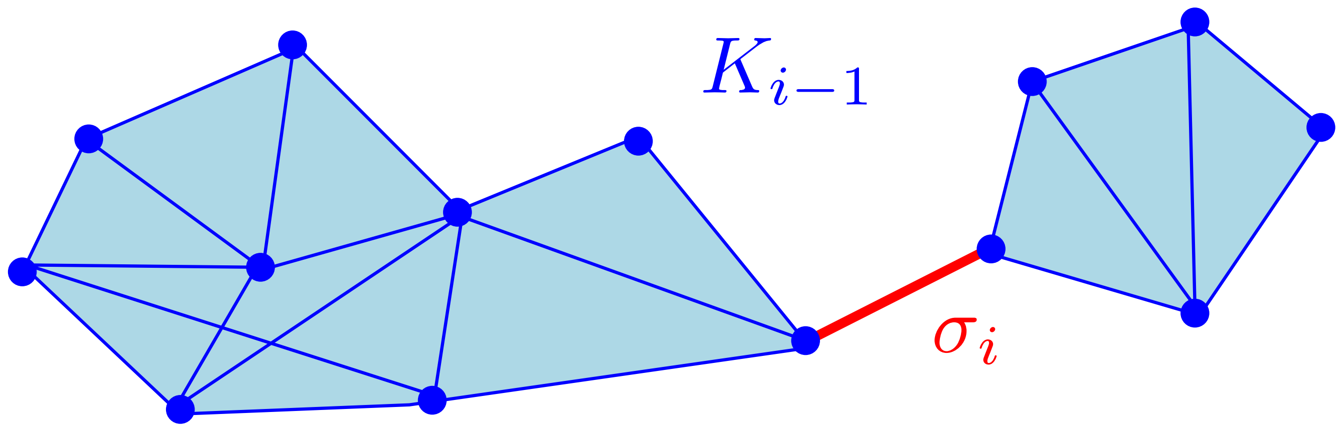
Let $k = \dim \sigma_i$ and denote $K_{i-1} = \cup_{l=1}^{i-1} \sigma_l$

Case 1 : adding σ_i to K_{i-1} creates a new k-dimensional topological feature in Ki (new homology class in H_k).



$\Rightarrow$ the birth of a k-dim feature is registered.

Case 2 : adding σ_i to K_{i-1} kills a $(k-1)$-dimensional topological feature in K_i (homology class in H_{k-1}).



$\Rightarrow$ persistence algo. pairs the simplex σ_i to the simplex $\sigma_{l(i)}$ that gave birth to the killed feature.

Key remark : the persistence pairs are determined by the order on the simplices ; the corresponding points in the diagrams are determined by the filtration values.

$\rightarrow$ $(\sigma_{l(i)}, \sigma_i)$: persistence pair

$\rightarrow$ $(\Phi_{\sigma_{l(i)}}, \Phi_{\sigma_i}) \in \mathbb{R}^2$: point in the persistence diagram

Persistent homology computation

Filtration

$\Phi = (\Phi_\sigma)_{\sigma \in K} \in \mathbb{R}^{|K|}$

Pairing of simplices

p pairs $(\sigma_{l(i)}, \sigma_i)$

q unpaired σ_l

$|K| = 2p + q$

Persistence diagram $D(\Phi)$

$(\Phi_{\sigma_{l(i)}}, \Phi_{\sigma_i})$

$(\Phi_{\sigma_l}, +\infty)$

Using lexicographical order

$D(\Phi) \in \mathbb{R}^{|K|}$

The persistence map : Pers

(this is a "locally constant" permutation of coordinates)

The persistence map is semi-algebraic

Proposition : Given a simplicial complex K, the map

$$\mathrm{Pers}\colon \mathrm{Filt}_K \subseteq \mathbb{R}^{|K|} \to \mathbb{R}^{|K|}$$

is semi-algebraic, and thus definable in any o-minimal structure. Moreover, there exists a semi-algebraic partition of Filt_K such that the restriction of Pers to each element of this partition is a Lipschitz map.

Corollary : Let K be a simplicial complex and $\Phi\colon A \to \mathbb{R}^{|K|}$ be a definable (in a given o-minimal structure) parametrized family of filtrations. The map $\mathrm{Pers} \circ \Phi\colon A \to \mathbb{R}^{|K|}$ is definable.

The persistence map is semi-algebraic

Proposition : Let K be a simplicial complex and $\Phi\colon A \to \mathbb{R}^{|K|}$ a definable parametrized family of filtrations, where $\dim A = m$. Then there exists a finite definable partition of A, $A = S \sqcup O_1 \sqcup \cdots \sqcup O_k$ such that $\dim S < \dim A := m$ and, for any $i = 1, \ldots, k$, O_i is a definable manifold of dimension m and $\mathrm{Pers} \circ \Phi\colon O_i \to \mathbb{R}^{|K|}$ is differentiable.

This is an immediate consequence of finiteness and stratifiability properties of definable sets

o-minimal structures

An o-minimal structure on the field of real numbers $\mathbb{R}$ is a collection $(S_n)_{n\in\mathbb{N}}$, where each S_n is a set of subsets of $\mathbb{R}^n$ such that :

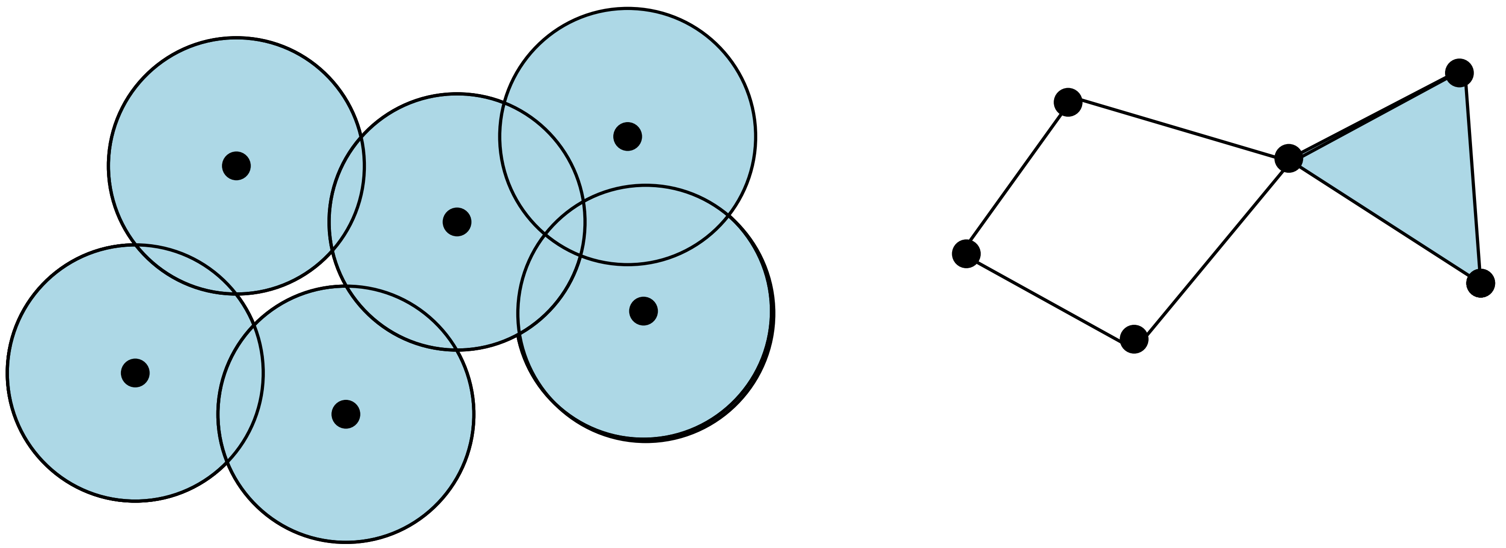
1. S_1 is exactly the collection of finite unions of points and intervals ;
2. all algebraic subsets of $\mathbb{R}^n$ are in S_n ;
3. S_n is a Boolean subalgebra of $\mathbb{R}^n$ for any $n \in \mathbb{N}$;
4. if $A \in S_n$ and $B \in S_m$, then $A \times B \in S_{n+m}$;
5. if $\pi\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the linear projection onto the first n coordinates and $A \in S_{n+1}$, then $\pi(A) \in S_n$.

$A \in S_n$ is called a definable set in the o-minimal structure.

For $A \subseteq \mathbb{R}^n$, a map $f\colon A \to \mathbb{R}^m$ is a definable map if its graph is a definable set in $\mathbb{R}^{n+m}$.

Important property : Definable sets admit finite (Whitney) stratification.

Example : the Vietoris-Rips filtration

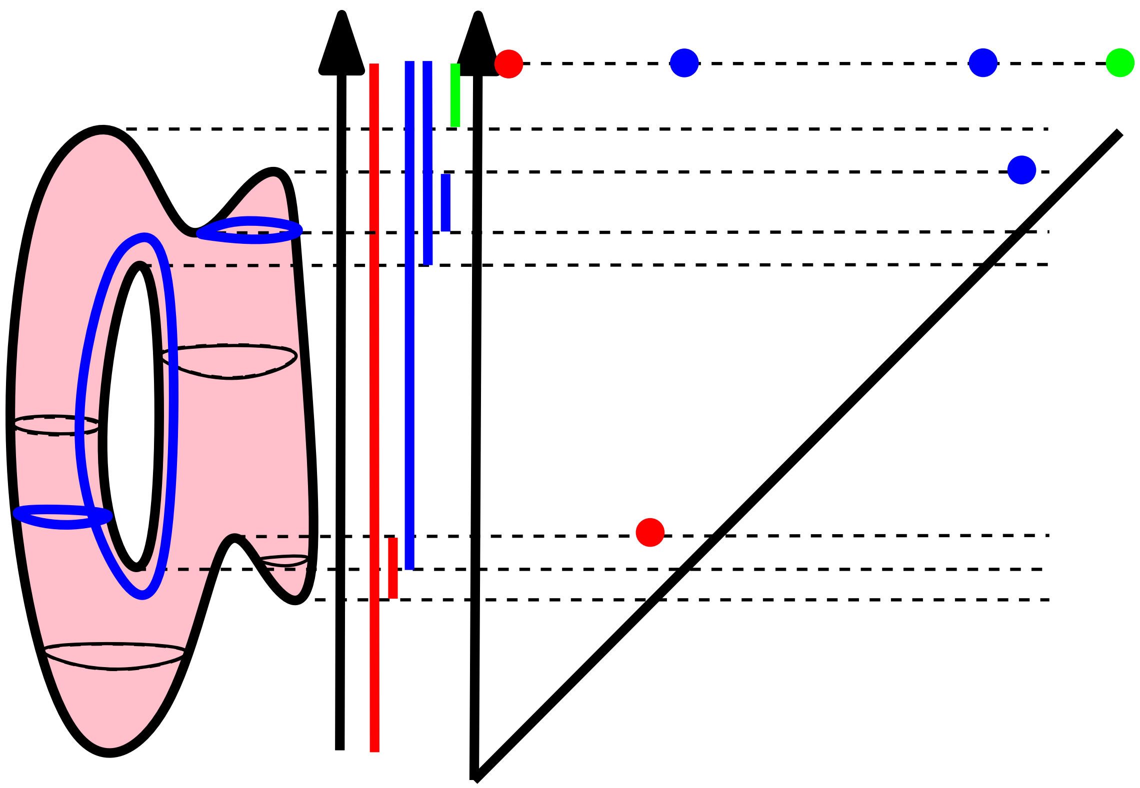


$$\Phi\colon A = (\mathbb{R}^d)^n \to \mathbb{R}^{|\Delta_n|} = \mathbb{R}^{2^n - 1}$$

where Δ_n is the simplicial complex made of all the faces of the $(n-1)$-dimensional simplex and, for any $x = (x_1, \ldots, x_n) \in A$ and any simplex $\sigma \subseteq \{1, \ldots, n\}$,

$$\Phi_\sigma(x) = \max_{i,j \in \sigma} \|x_i - x_j\|.$$

Example : sublevel sets filtrations



K a simplicial complex with n vertices $v_1, \ldots, v_n$.

Any real-valued function f defined on the vertices of K can be represented as a vector $(f(v_1), \ldots, f(v_n)) \in \mathbb{R}^n$.

$$\Phi\colon A = \mathbb{R}^n \to \mathbb{R}^{|K|}$$

where for any $f = (f_1, \ldots, f_n) \in A$ and any simplex $\sigma \subseteq \{1, \ldots, n\}$,

$$\Phi_\sigma(f) = \max_{i \in \sigma} f_i$$

Functions of persistence

Definition : A function

$$E\colon \mathbb{R}^{|K|} = (\mathbb{R}^2)^p \times \mathbb{R}^q \to \mathbb{R}$$

is a function of persistence if it is invariant to permutations of the points of the persistence diagram : for any $(p_1, \ldots, p_p, e_1, \ldots, e_q) \in (\mathbb{R}^2)^p \times \mathbb{R}^q$ and any permutations α, β of the sets $\{1, \ldots, p\}$ and $\{1, \ldots, q\}$, respectively, one has

$$E(p_{\alpha(1)}, \ldots, p_{\alpha(p)}, e_{\beta(1)}, \ldots, e_{\beta(q)}) = E(p_1, \ldots, p_p, e_1, \ldots, e_q).$$

Functions of persistence

Definition : A function

$$E\colon \mathbb{R}^{|K|} = (\mathbb{R}^2)^p \times \mathbb{R}^q \to \mathbb{R}$$

is a function of persistence if it is invariant to permutations of the points of the persistence diagram : for any $(p_1, \ldots, p_p, e_1, \ldots, e_q) \in (\mathbb{R}^2)^p \times \mathbb{R}^q$ and any permutations α, β of the sets $\{1, \ldots, p\}$ and $\{1, \ldots, q\}$, respectively, one has

$$E(p_{\alpha(1)}, \ldots, p_{\alpha(p)}, e_{\beta(1)}, \ldots, e_{\beta(q)}) = E(p_1, \ldots, p_p, e_1, \ldots, e_q).$$

Properties :

If E is locally Lipschitz, then the composition $E \circ \mathrm{Pers}$ is also locally Lipschitz.

If E and $\Phi\colon A \subseteq \mathbb{R}^d \to \mathbb{R}^{|K|}$ are definable, then $\mathcal{L} = E \circ \mathrm{Pers} \circ \Phi\colon A \to \mathbb{R}$ has a well-defined Clarke subdifferential $\partial\mathcal{L}(z) := \mathrm{Conv}\{\lim_{z_i \to z} \nabla\mathcal{L}(z_i) \; : \; \mathcal{L}$ is differentiable at $z_i\}$.

Examples

Total persistence.

$$E(D) = \sum_{i=1}^{p} |d_i - b_i|, \quad \text{for} \quad D = ((b_1, d_1), \ldots, (b_p, d_p), e_1, \ldots, e_q).$$

E is semi-algebraic and Lipschitz.

Examples

Total persistence.

$$E(D) = \sum_{i=1}^{p} |d_i - b_i|, \quad \text{for} \quad D = ((b_1, d_1), \ldots, (b_p, d_p), e_1, \ldots, e_q).$$

E is semi-algebraic and Lipschitz.

Bottleneck distance.

$$E(D) = d_B(D, D^*) = \min_{m} \max_{(p,p^*)\in m} ||p - p^*||_\infty$$

where denoting $\Delta = \{(x, x) : x \in \mathbb{R}\}$ the diagonal in $\mathbb{R}^2$, m is a partial matching between D and D^*, i.e., a subset of $(D \cup \Delta) \times (D^* \cup \Delta)$ such that every point of $D \setminus \Delta$ and $D^* \setminus \Delta$, appears exactly once in m.
E is semi-algebraic and Lipschitz.

Minimization via stochastic (sub-)gradient descent

If E and $\Phi\colon A \subseteq \mathbb{R}^d \to \mathbb{R}^{|K|}$ are definable, then $\mathcal{L} = E \circ \mathrm{Pers} \circ \Phi\colon A \to \mathbb{R}$ has a well-defined Clarke subdifferential $\partial\mathcal{L}(z) := \mathrm{Conv}\{\lim_{z_i \to z} \nabla\mathcal{L}(z_i) \ : \ \mathcal{L}$ is differentiable at $z_i\}$.

Minimization of $\mathcal{L}$ through the differential inclusion

$$\frac{dz}{dt} \in -\partial\mathcal{L}(z(t)) \quad \text{for almost every } t.$$

Standard stochastic subgradient algorithm

$$x_{k+1} = x_k - \alpha_k(y_k + \zeta_k), \ y_k \in \partial\mathcal{L}(x_k),$$

where the sequence $(\alpha_k)_k$ is the learning rate and $(\zeta_k)_k$ is a sequence of random variables.

Minimization via stochastic (sub-)gradient descent

If E and $\Phi\colon A \subseteq \mathbb{R}^d \to \mathbb{R}^{|K|}$ are definable, then $\mathcal{L} = E \circ \mathrm{Pers} \circ \Phi\colon A \to \mathbb{R}$ has a well-defined Clarke subdifferential $\partial\mathcal{L}(z) := \mathrm{Conv}\{\lim_{z_i \to z} \nabla\mathcal{L}(z_i) \;:\; \mathcal{L}$ is differentiable at $z_i\}$.

Minimization of $\mathcal{L}$ through the differential inclusion

$$\frac{dz}{dt} \in -\partial\mathcal{L}(z(t)) \quad \text{for almost every } t.$$

Standard stochastic subgradient algorithm

$$x_{k+1} = x_k - \alpha_k(y_k + \zeta_k),\; y_k \in \partial\mathcal{L}(x_k),$$

where the sequence $(\alpha_k)_k$ is the learning rate and $(\zeta_k)_k$ is a sequence of random variables.

Question : convergence of the algorithm ?

Convergence

Convergence follows from [Davis et al, Stochastic subgradient method converges on tame functions. Found. Comp. Math. 2020].

Standard stochastic subgradient algorithm

$$x_{k+1} = x_k - \alpha_k(y_k + \zeta_k),\ y_k \in \partial \mathcal{L}(x_k),$$

where the sequence $(\alpha_k)_k$ is the learning rate and $(\zeta_k)_k$ is a sequence of random variables.

Technical (but classical) assumptions :

1. for any k, $\alpha_k \geq 0$, $\sum_{k=1}^{\infty} \alpha_k = +\infty$ and, $\sum_{k=1}^{\infty} \alpha_k^2 < +\infty$;
2. $\sup_k \|x_k\| < +\infty$, almost surely ;
3. denoting by $\mathcal{F}_k$ the increasing sequence of σ-algebras $\mathcal{F}_k = \sigma(x_j, y_j, \zeta_j, j < k)$, there exists a function $p \colon \mathbb{R}^d \to \mathbb{R}$ which is bounded on bounded sets such that almost surely, for any k,

$$\mathbb{E}[\zeta_k|\mathcal{F}_k] = 0 \text{ and } \mathbb{E}[\|\zeta_k\|^2|\mathcal{F}_k] < p(x_k).$$

Convergence

Convergence follows from [Davis et al, Stochastic subgradient method converges on tame functions. Found. Comp. Math. 2020].

Standard stochastic subgradient algorithm

$$x_{k+1} = x_k - \alpha_k(y_k + \zeta_k),\ y_k \in \partial \mathcal{L}(x_k),$$

where the sequence $(\alpha_k)_k$ is the learning rate and $(\zeta_k)_k$ is a sequence of random variables.

Theorem :
Let K be a simplicial complex, $A \subseteq \mathbb{R}^d$, and $\Phi\colon A \to \mathbb{R}^{|K|}$ a parametrized family of filtrations of K that is definable in an o-minimal structure. Let $E\colon \mathbb{R}^{|K|} \to \mathbb{R}$ be a definable function of persistence such that $\mathcal{L} = E \circ \mathrm{Pers} \circ \Phi$ is locally Lipschitz. Then, under the above assumptions 1, 2, and 3, almost surely the limit points of the sequence $(x_k)_k$ obtained from the iterations of the algo. are critical points of $\mathcal{L}$ and the sequence $(\mathcal{L}(x_k))_k$ converges.

Numerical illustration

The differential of persistence map is obvious to compute $\rightarrow$ easy implementation (soon available in GUDHI)

Point cloud optimization

Input : a point cloud X sampled uniformly from the unit square $S = [0,1]^2$

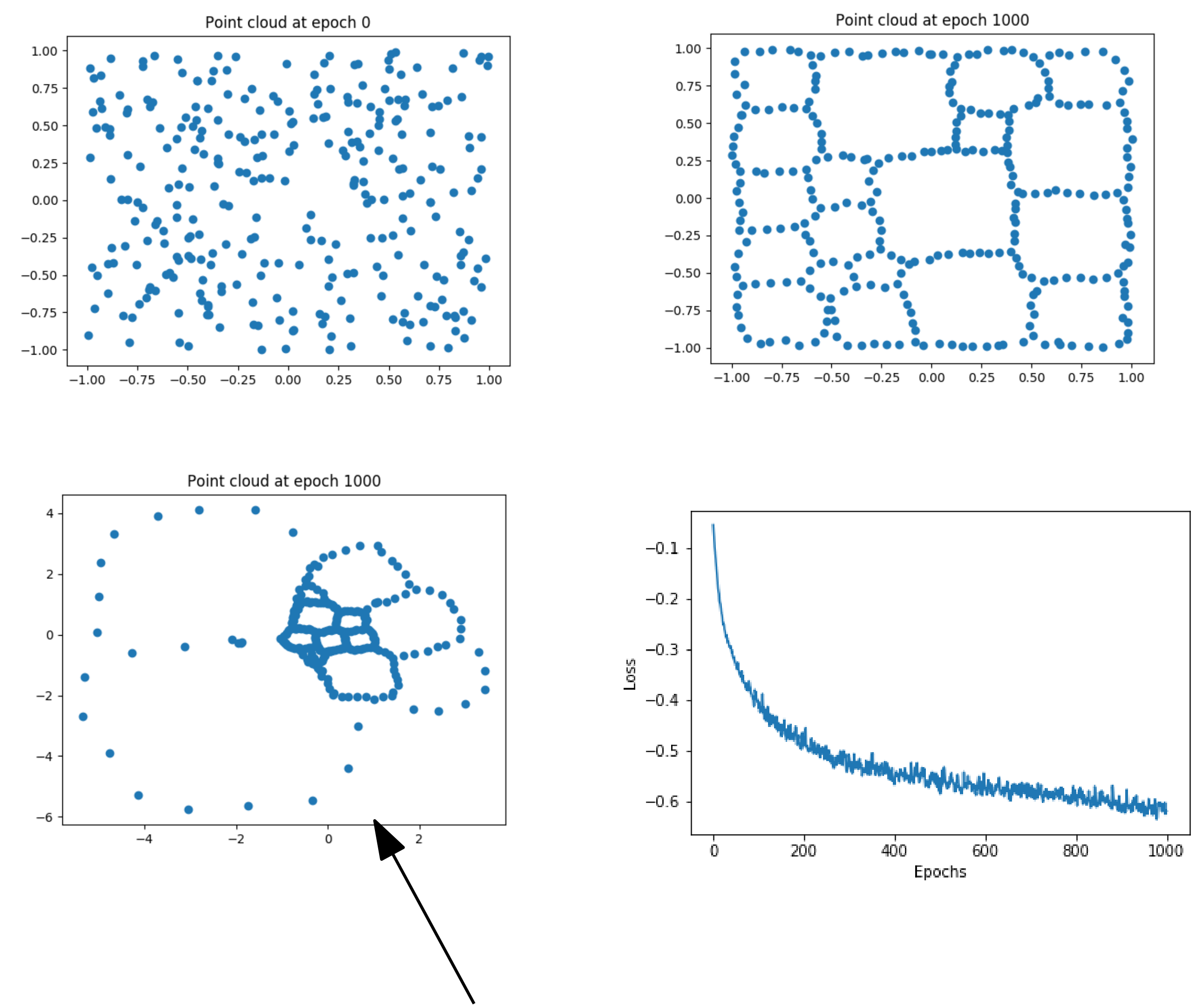
Loss : $\mathcal{L}(X) = P(X) + T(X)$ where

$$T(X) := -\sum_{p \in D} \|p - \pi_\Delta(p)\|_\infty^2$$

with D is the 1-dimensional persistence diagram associated to the Vietoris-Rips filtration of X, π_Δ stands for the projection onto the diagonal Δ, and

$$P(X) := \sum_{x \in X} d(x, S)$$

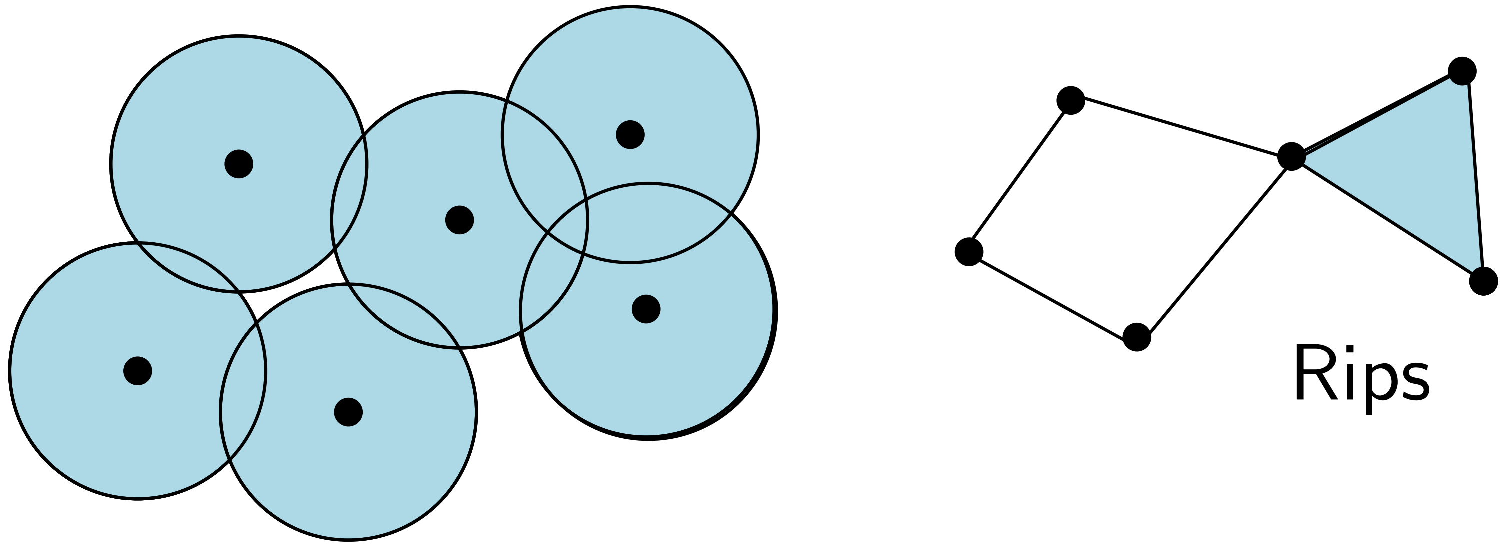
is a penalty term ensuring that the point coordinates stay in the unit square.



With $T(X)$ only

The density of expected persistence diagrams
(The Vietoris-Rips case)

The Vietoris-Rips filtration



Let V be a point cloud (in a metric space (X, d)).

The Vietoris-Rips complex $\mathrm{Rips}(V)$ is the filtered simplicial complex indexed by $\mathbb{R}$ whose vertex set is V and defined by :

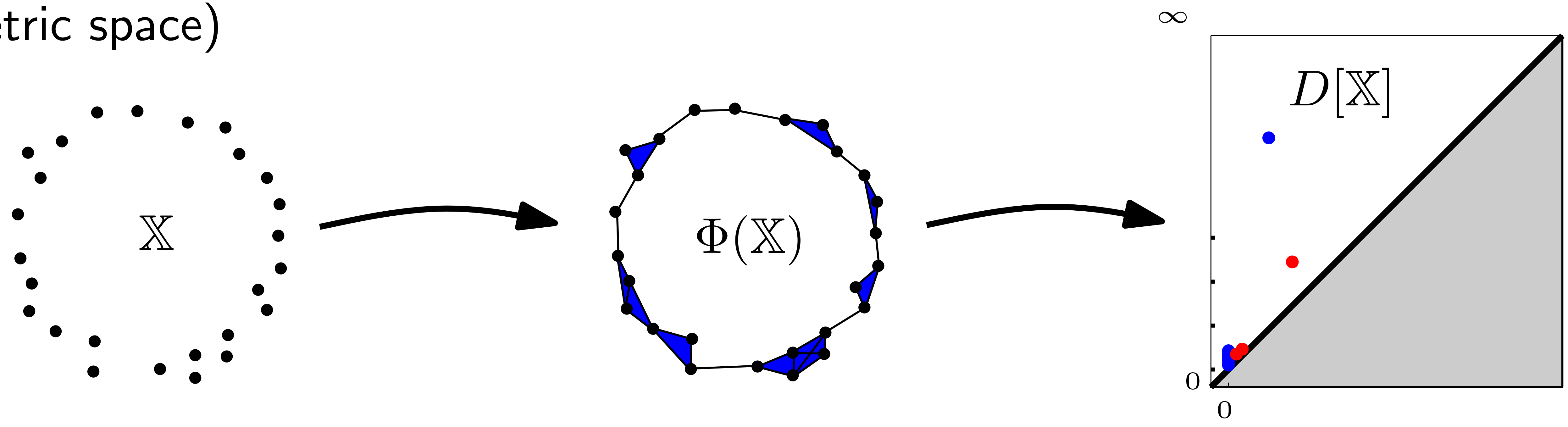
$$\Phi_\sigma(V) = \max_{v,v' \in \sigma} d(v, v')$$

Statistical setting

$\mathbb{X}$ is now a random point coud (in some metric space)

Φ is the (deterministic) VR filtration

$D[\mathbb{X}]$ becomes random

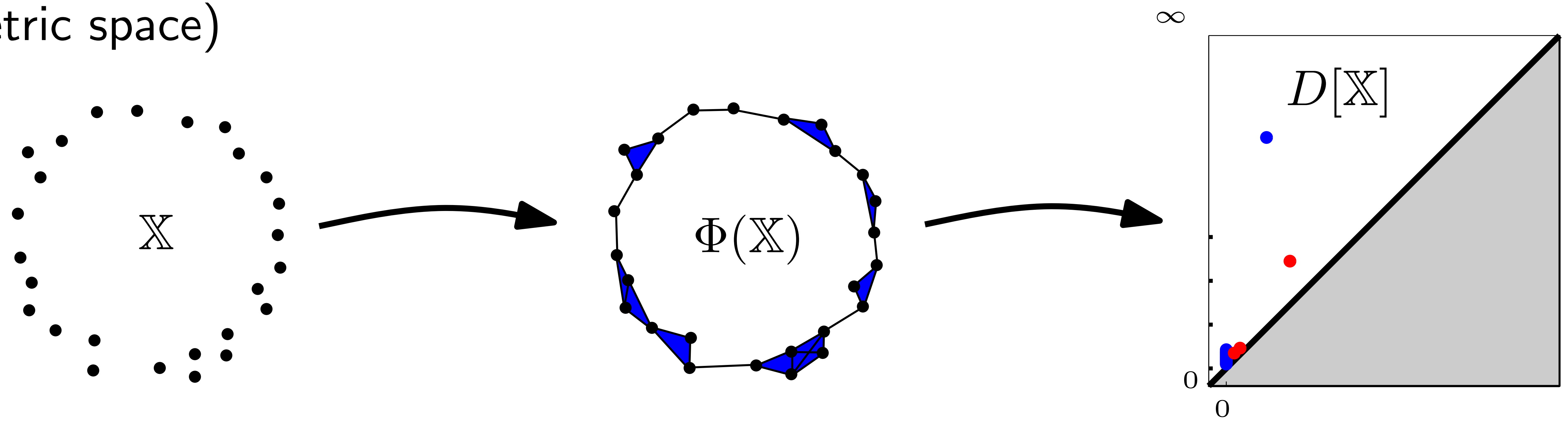


Statistical setting

$\mathbb{X}$ is now a random point coud (in some metric space)

Φ is the (deterministic) VR filtration

$D[\mathbb{X}]$ becomes random



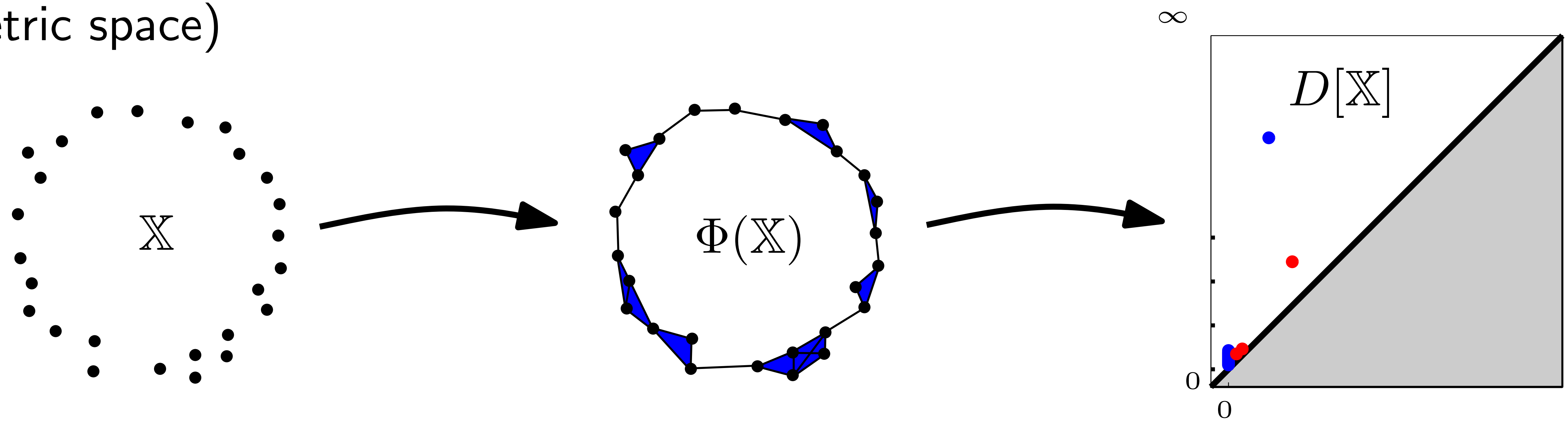
What can be said about the distribution of diagrams $D[\mathbb{X}]$?

Statistical setting

$\mathbb{X}$ is now a random point coud (in some metric space)

Φ is the (determi-nistic) VR filtration

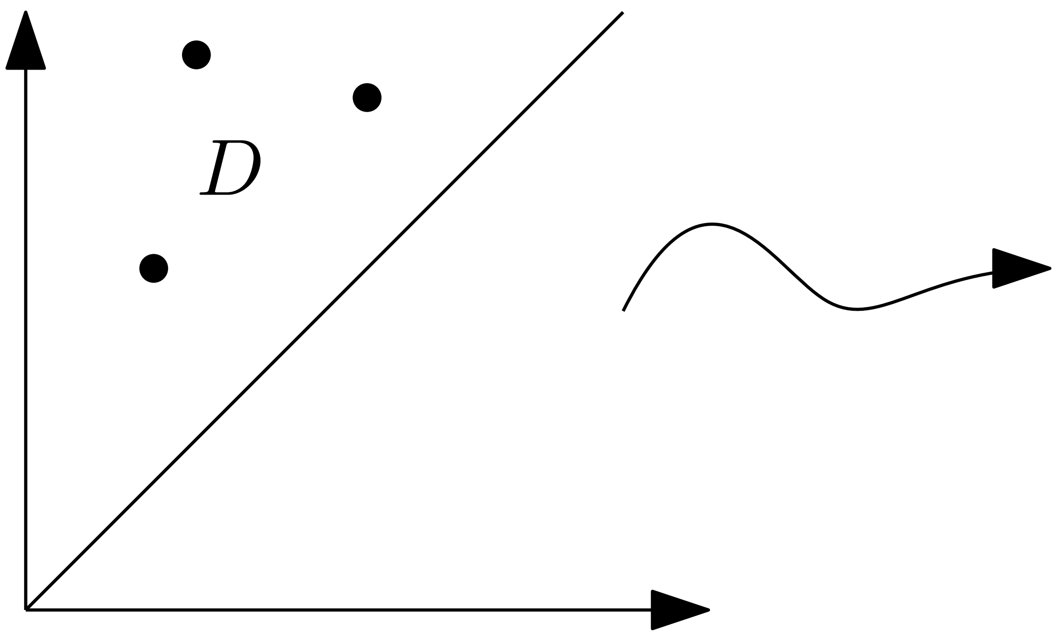
$D[\mathbb{X}]$ becomes random



What can be said about the distribution of diagrams $D[\mathbb{X}]$?

Understand the structure of $E[D[\mathbb{X}]]$ in the non asymptotic setting ($|\mathbb{X}| = n$ is fixed, or bounded)

Persistence diagrams as discrete measures



$$D := \sum_{p \in D} \delta_p$$

Motivations :

- The space of measures is much nicer that the space of P. D. !
- In the general algebraic persistence theory, persistence diagrams naturally appears as discrete measures in the plane.

[C., de Silva, Glisse, Oudot 16]

- Many persistence representations can be expressed as

$$D(f) = \sum_{p \in D} f(p) = \int f dD$$

for well-chosen functions $f : \mathbb{R}^2 \to \mathcal{H}$.

The density of expected persistence diagrams

Theorem :
Fix $n \geq 1$. Assume that :

- M is a real analytic compact d-dimensional connected riemannian manifold possibly with boundary,
- $\mathbb{X}$ is a random variable on M^n having a density with respect to the Haussdorf measure $\mathcal{H}_{dn}$,
- Φ is the Vietoris-Rips filtration and denote $D_s[\Phi]$ its s-dimensional persistence diagram.

Then, for $s \geq 1$, $E[D_s[\Phi(\mathbb{X})]]$ has a density with respect to the Lebesgue measure on Δ, the upper half-plane above the diagonal. Moreover, $E[D_0[\Phi(\mathbb{X})]]$ has a density with respect to the Lebesgue measure on the vertical line $\{0\} \times [0, \infty)$.

The density of expected persistence diagrams

Theorem :
Fix $n \geq 1$. Assume that :

- M is a real analytic compact d-dimensional connected riemannian manifold possibly with boundary,
- $\mathbb{X}$ is a random variable on M^n having a density with respect to the Haussdorf measure $\mathcal{H}_{dn}$,
- Φ is the Vietoris-Rips filtration and denote $D_s[\Phi]$ its s-dimensional persistence diagram.

Then, for $s \geq 1$, $E[D_s[\Phi(\mathbb{X})]]$ has a density with respect to the Lebesgue measure on Δ, the upper half-plane above the diagonal. Moreover, $E[D_0[\Phi(\mathbb{X})]]$ has a density with respect to the Lebesgue measure on the vertical line $\{0\} \times [0, \infty)$.

Theorem [smoothness] : Under the assumption of previous theorem, if moreover $\mathbb{X} \in M^n$ has a density of class C^k with respect to $\mathcal{H}_{nd}$. Then, for $s \geq 0$, the density of $E[D_s[\Phi(\mathbb{X})]]$ is of class C^k.

Sketch of proof ($s \geq 1$)

1. There exists a partition of the complement of a (subanalytic) set of measure 0 in M^n by open sets $V_1, \cdots, V_R$ such that :

- the order of the simplices of $\Phi(x)$ is constant on each V_r,
- for any $r = 1, \cdots, R$, and any $x \in V_r$,

$$D_s[\Phi(x)] = \sum_{i=1}^{N_r} \delta_{\mathbf{r}_i}$$

with $\mathbf{r}_i = (\Phi_{\sigma_{i_1}}(x), \Phi_{\sigma_{i_2}}(x))$ where N_r, $\sigma_{i_1}, \sigma_{i_2}$ only depends on V_r.

- $\sigma_{i_1}, \sigma_{i_2}$ can be chosen so that the differential of

$$\Phi_{ir} : x \in V_r \to \mathbf{r}_i = (\Phi_{\sigma_{i_1}}(x), \Phi_{\sigma_{i_2}}(x))$$

has maximal rank 2.

Sketch of proof ($s \geq 1$)

2. The expected diagram can be written as

$$
\begin{aligned}
E[D_s[\Phi(\mathbb{X})]] &= \sum_{r=1}^{R} E\left[\mathbb{1}\{\mathbb{X} \in V_r\} D_s[\Phi(\mathbb{X})]\right] = \sum_{r=1}^{R} E\left[\mathbb{1}\{\mathbb{X} \in V_r\} \sum_{i=1}^{N_r} \delta_{\mathbf{r}_i}\right] \\
&= \sum_{r=1}^{R} \sum_{i=1}^{N_r} E\left[\mathbb{1}\{\mathbb{X} \in V_r\} \delta_{\mathbf{r}_i}\right]
\end{aligned}
$$

Sketch of proof ($s \geq 1$)

2. The expected diagram can be written as

$$E[D_s[\Phi(\mathbb{X})]] = \sum_{r=1}^{R} E\left[\mathbb{1}\{\mathbb{X} \in V_r\} D_s[\Phi(\mathbb{X})]\right] = \sum_{r=1}^{R} E\left[\mathbb{1}\{\mathbb{X} \in V_r\} \sum_{i=1}^{N_r} \delta_{\mathbf{r}_i}\right]$$

$$= \sum_{r=1}^{R} \sum_{i=1}^{N_r} E\left[\mathbb{1}\{\mathbb{X} \in V_r\} \delta_{\mathbf{r}_i}\right]$$

μ_{ir}

3. Use the co-area formula :

$$\mu_{ir}(B) = P(\Phi_{ir}(\mathbb{X}) \in B, \mathbb{X} \in V_r)$$

$$= \int_{V_r} \mathbb{1}\{\Phi_{ir}(x) \in B\} \kappa(x) d\mathcal{H}_{nd}(x)$$

$$= \int_{u \in B} \int_{x \in \Phi_{ir}^{-1}(u)} (J\Phi_{ir}(x))^{-1} \kappa(x) d\mathcal{H}_{nd-2}(x) du.$$

Density of $\mathbb{X}$

Density of μ_{ir}

The Hausdorff measure and the co-area formula

Definition : Let k be a non-negative number. For $A \subset \mathbb{R}^D$, and $\delta > 0$, consider

$$\mathcal{H}_k^\delta(A) := \inf\left\{\sum_i \text{diam}(U_i)^k, A \subset \bigcup_i U_i \text{ and } \text{diam}(U_i) < \delta\right\}.$$

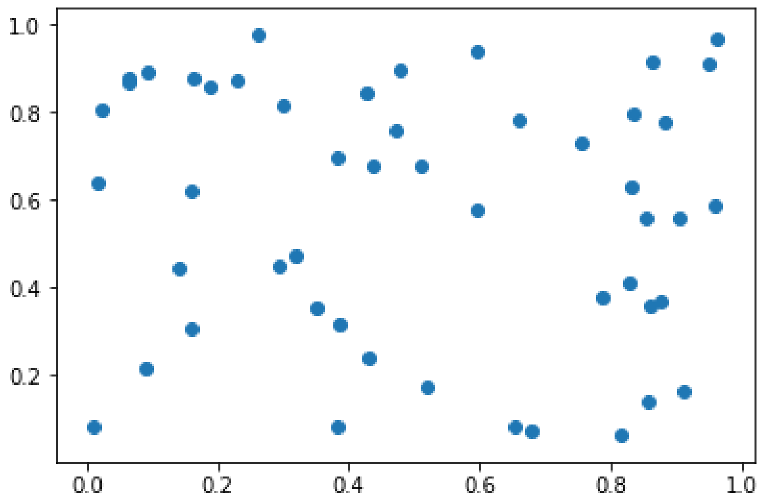
The *k-dimensional Haussdorf measure* on $\mathbb{R}^D$ of A is defined by $\mathcal{H}_k(A) := \lim_{\delta \to 0} \mathcal{H}_k^\delta(A)$.

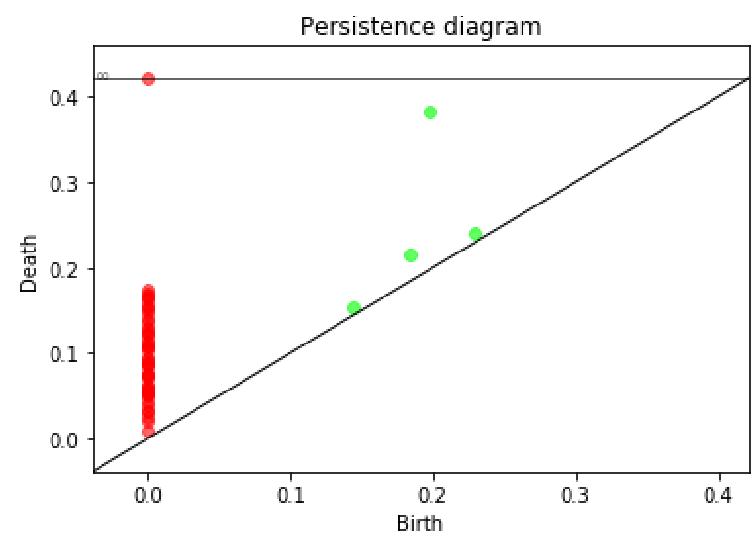
Theorem [Co-area formula] : Let M (resp. N) be a smooth Riemannian manifold of dimension m (resp n). Assume that $m \geq n$ and let $\Phi : M \to N$ be a differentiable map. Denote by $D\Phi$ the differential of Φ. The Jacobian of Φ is defined by $J\Phi = \sqrt{\det((D\Phi) \times (D\Phi)^t)}$. For $f : M \to N$ a positive measurable function, the following equality holds :

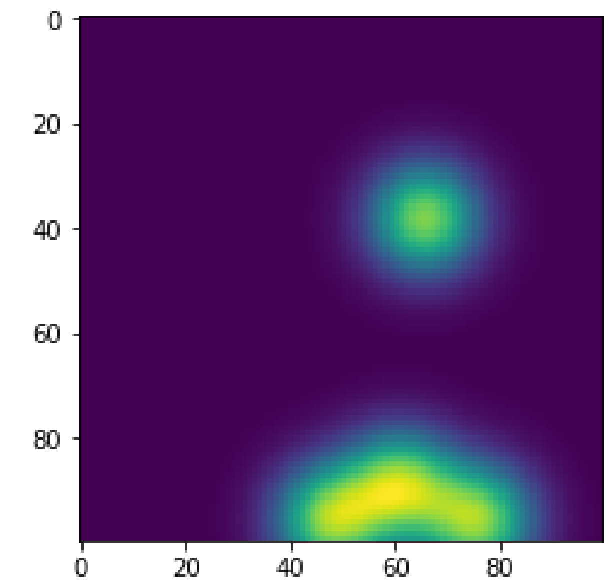
$$\int_M f(x) J\Phi(x) d\mathcal{H}_m(x) = \int_N \left(\int_{x \in \Phi^{-1}(\{y\})} f(x) d\mathcal{H}_{m-n}(x)\right) d\mathcal{H}_n(y).$$

Persistence images

[Adams et al, JMLR 2017]







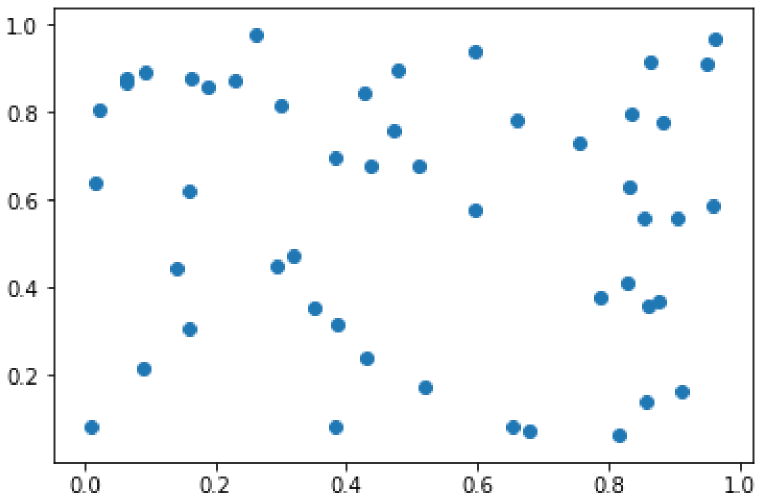
For $K : \mathbb{R}^2 \to \mathbb{R}$ a kernel and H a bandwidth matrix (e.g. a symmetric positive definite matrix), pose for $u \in \mathbb{R}^2$, $K_H(z) = |H|^{-1/2} K(H^{-1/2} \cdot u)$

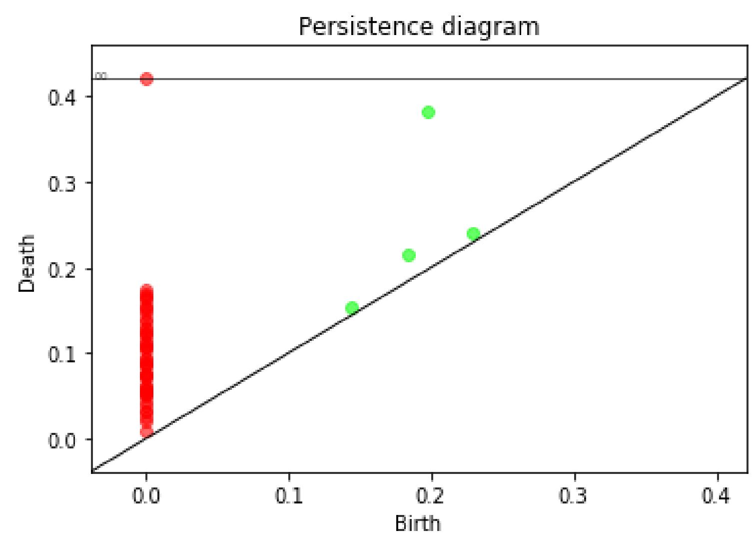
For $D = \sum_i \delta_{\mathbf{r}_i}$ a diagram, $K : \mathbb{R}^2 \to \mathbb{R}$ a kernel, H a bandwidth matrix and $w : \mathbb{R}^2 \to \mathbb{R}_+$ a weight function, one defines the persistence surface of D with kernel K and weight function w by :

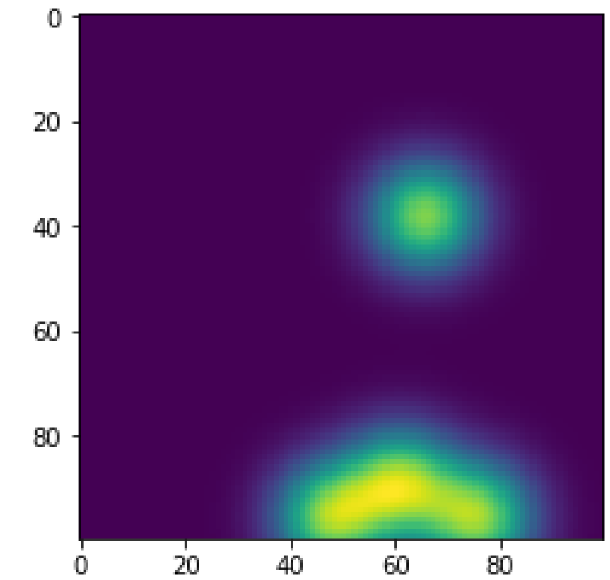
$$\forall z \in \mathbb{R}^2,\ \rho(D)(u) = \sum_i w(\mathbf{r}_i) K_H(u - \mathbf{r}_i) = D(w K_H(u - \cdot))$$

Persistence images

[Adams et al, JMLR 2017]







For $K : \mathbb{R}^2 \to \mathbb{R}$ a kernel and H a bandwidth matrix (e.g. a symmetric positive definite matrix), pose for $u \in \mathbb{R}^2$, $K_H(z) = |H|^{-1/2} K(H^{-1/2} \cdot u)$

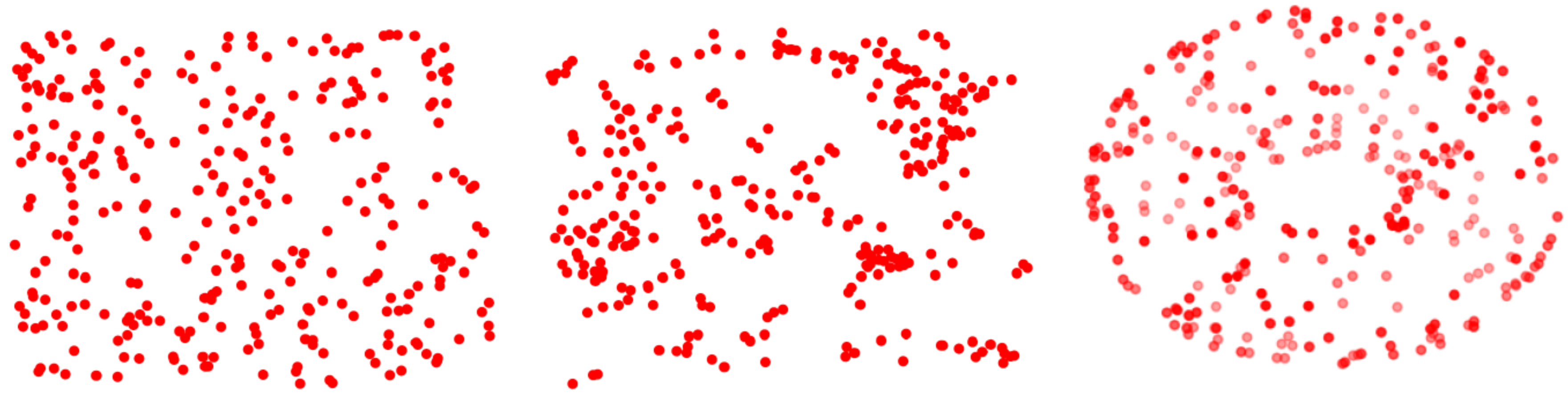
For $D = \sum_i \delta_{\mathbf{r}_i}$ a diagram, $K : \mathbb{R}^2 \to \mathbb{R}$ a kernel, H a bandwidth matrix and $w : \mathbb{R}^2 \to \mathbb{R}_+$ a weight function, one defines the persistence surface of D with kernel K and weight function w by :

$$\forall z \in \mathbb{R}^2,\ \rho(D)(u) = \sum_i w(\mathbf{r}_i) K_H(u - \mathbf{r}_i) = D(w K_H(u - \cdot))$$

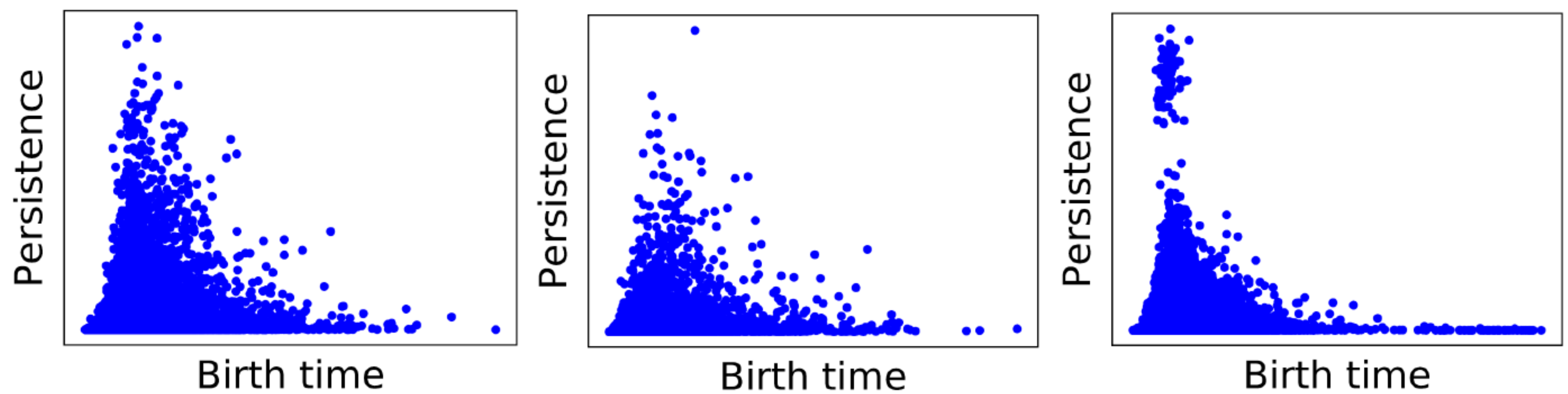
$\Rightarrow$ persistence surfaces can be seen as kernel estimates of $E[D_s[\Phi(\mathbb{X})]]$.

Persistence images

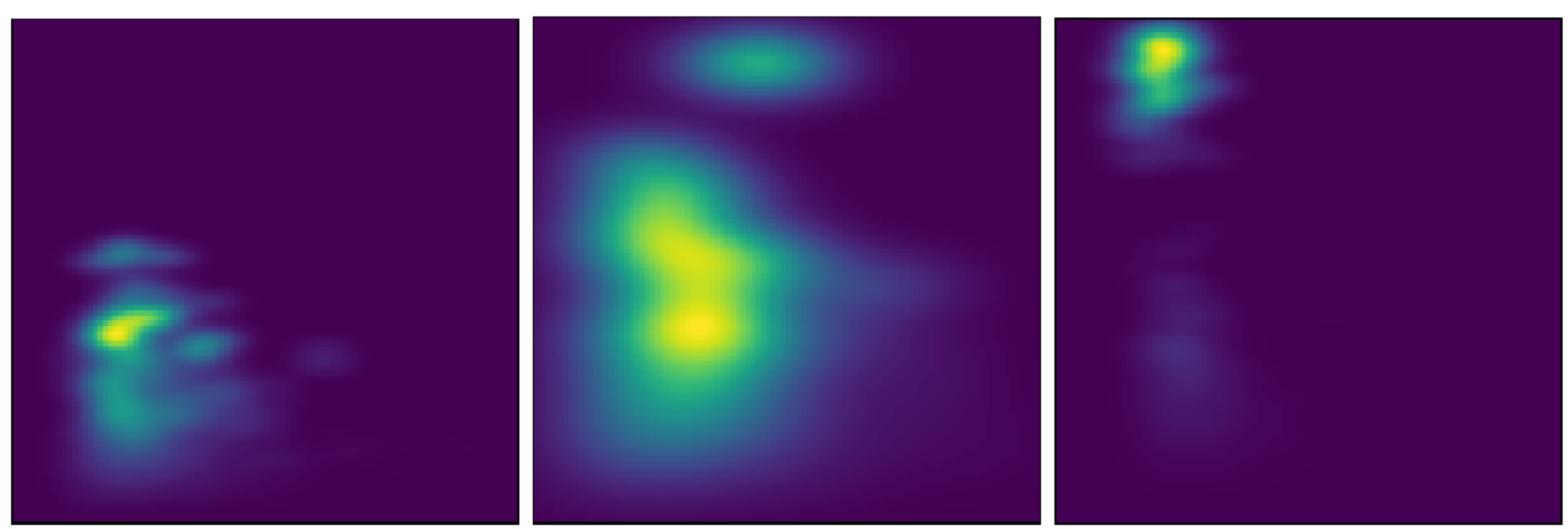
The realization of 3 different processes



The overlay of 40 different persistence diagrams



The persistence images with weight function $w(\mathbf{r}) = (r_2 - r_1)^3$ and bandwith selected using cross-validation.



Thank you for your attention !

References :

- M. Carriere, F. Chazal, M. Glisse, Y. Ike, H. Kannan. Optimizing persistent homology based functions. In Proc. ICML 2021
- V. Divol, F. Chazal. The density of expected persistence diagrams and its kernel based estimation. Journal of Computational Geometry 2019.

Persistence and TDA in practice :

- GUDHI library C++ / Python : https ://gudhi.inria.fr/