20/12/2024 Nice

Defense of the thesis

Persistent Geometry

by Antoine Commaret

Supervised by David COHEN-STEINER & Indira CHATTERJI

Defense of the thesis

Persistent Geometry

Presenting two results:

Persistent intrinsic volumes

Morse theory on tubular neighborhoods

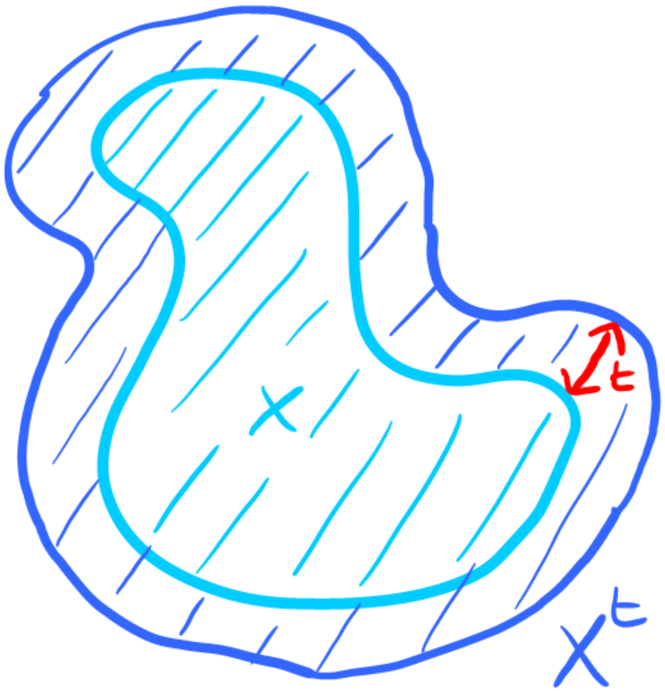
Objective

Let $X, Y \subset \mathbb{R}^d$

How can one recover the geometry of X from the knowledge of Y assuming X and Y are close?

Objective

Let $X, Y \subset \mathbb{R}^d$



Distance to X

$$X^t := \{x \in \mathbb{R}^d, d_X(x) \leq t\}$$

is the t-offset of X.

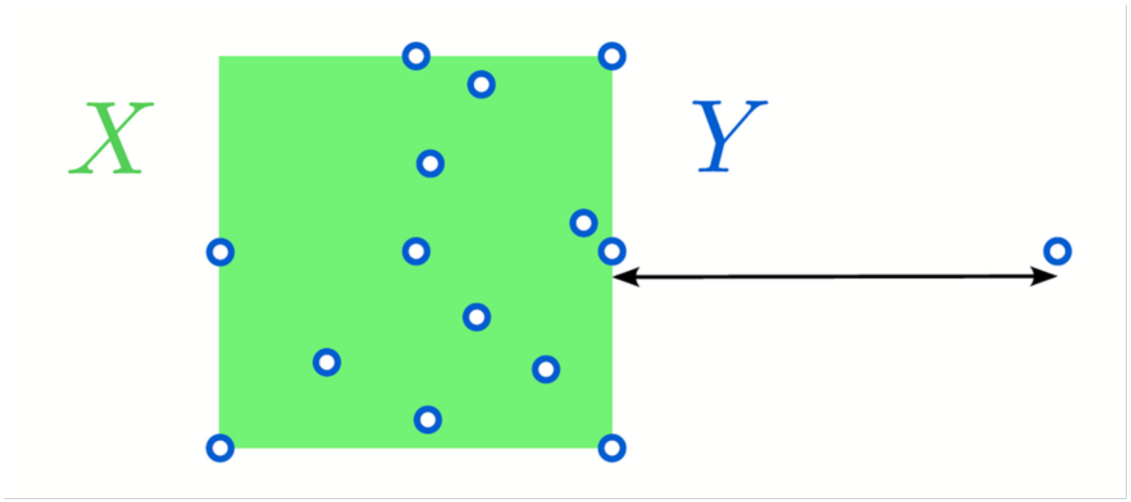
Hausdorff distance.

Let $X, Y \subset \mathbb{R}^d$

We use the Hausdorff distance

between X and Y

$$d_H(X,Y) = \inf\left\{ t \geq 0,\ X \subset Y^t,\ Y \subset X^t \right\}$$



Objective

$\Rightarrow$ We focus on the recovery of **Intrinsic volumes**

Intrinsic volumes

$\Rightarrow$ We focus on the recovery of <u>Intrinsic volumes</u>

Quantities $V_0(X), \ldots, V_{d-1}(X), V_d(X)$ associated to a <u>large class of sets in $\mathbb{R}^d$</u>.

Intrinsic volumes

$\Rightarrow$ We focus on the recovery of

Intrinsic volumes

Quantities $V_0(X), \ldots, V_{d-1}(X), V_d(X)$

Euler Characteristic $\chi(X)$

Boundary Area* $\mathcal{H}^{d-1}(\partial X)$

Volume $\mathcal{H}^d(X)$

Intrinsic volumes

Simple definition of intrinsic volumes via the Tube formula for sets with positive reach.

$$\mathrm{reach}(X) := \sup \left\{ t \in \mathbb{R} ,\ d_X(x) \leq t \Rightarrow x \text{ has a unique closest point in } X \right\}$$

Intrinsic volumes

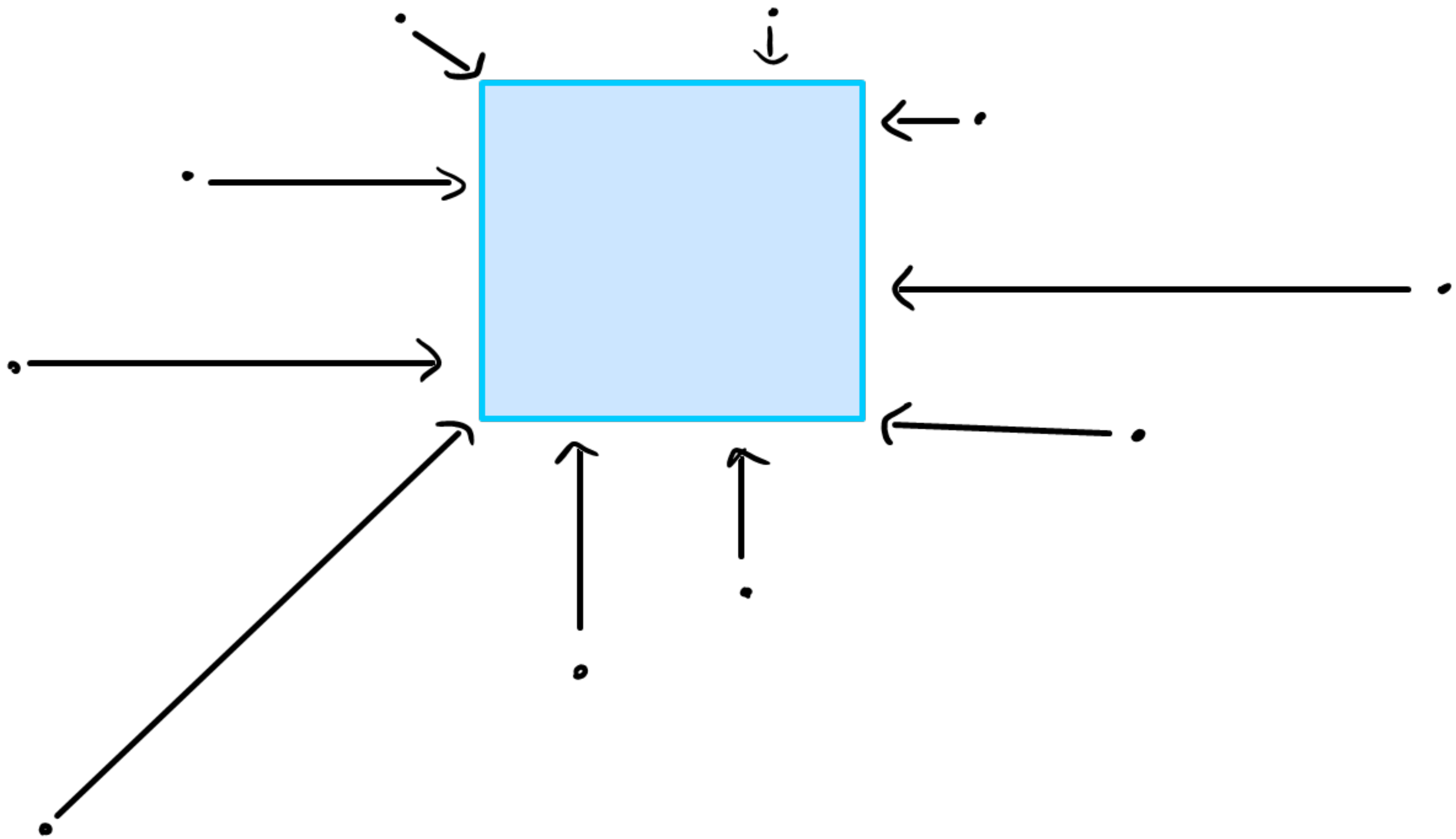
Simple definition of intrinsic volumes via the Tube formula for sets with positive reach.

$$\mathrm{reach}(X) := \sup \left\{ t \in \mathbb{R},\ d_X(x) \leq t \Rightarrow x \text{ has a unique closest point in } X \right\}$$

"the largest distance to X under which a point has a unique closest point in X"

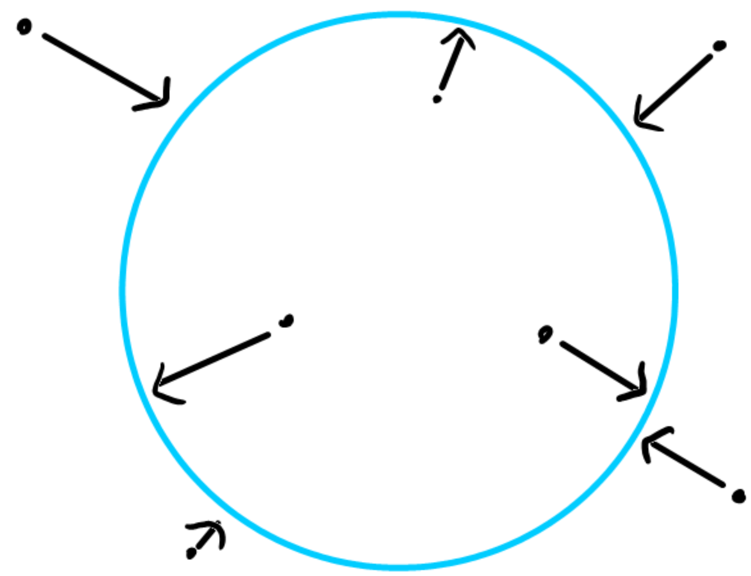
Examples:

A convex set has reach $+\infty$:



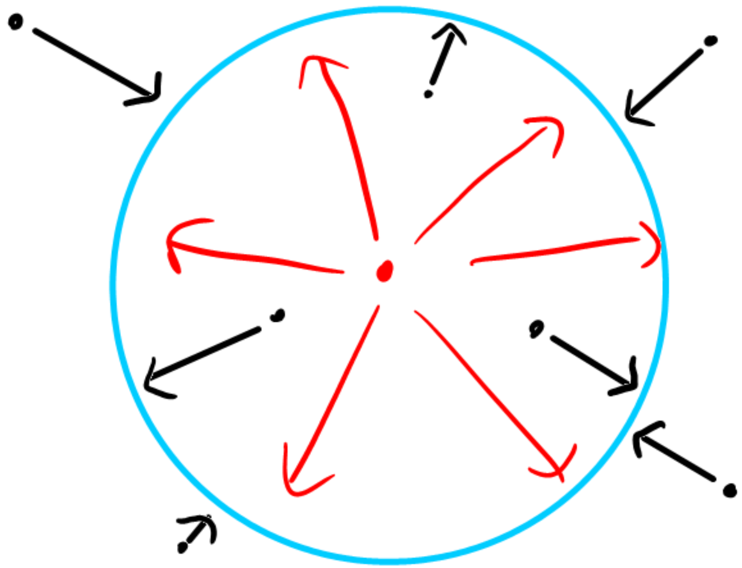
Examples:

Any compact submanifold of $\mathbb{R}^d$ has reach > 0.



Examples :

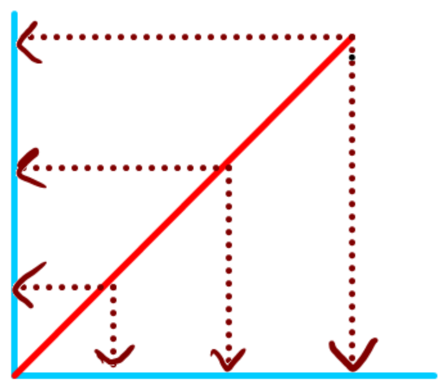
Any compact submanifold of $\mathbb{R}^d$ has reach > 0.



but no reason to be $+\infty$.

Examples :

Shapes with concave corners have reach zero.



Tube formula.

Within $[0, \mathrm{reach}(X)]$, (Federer 1959)

$t \longmapsto \mathrm{Vol}(X^t)$ is a polynomial.

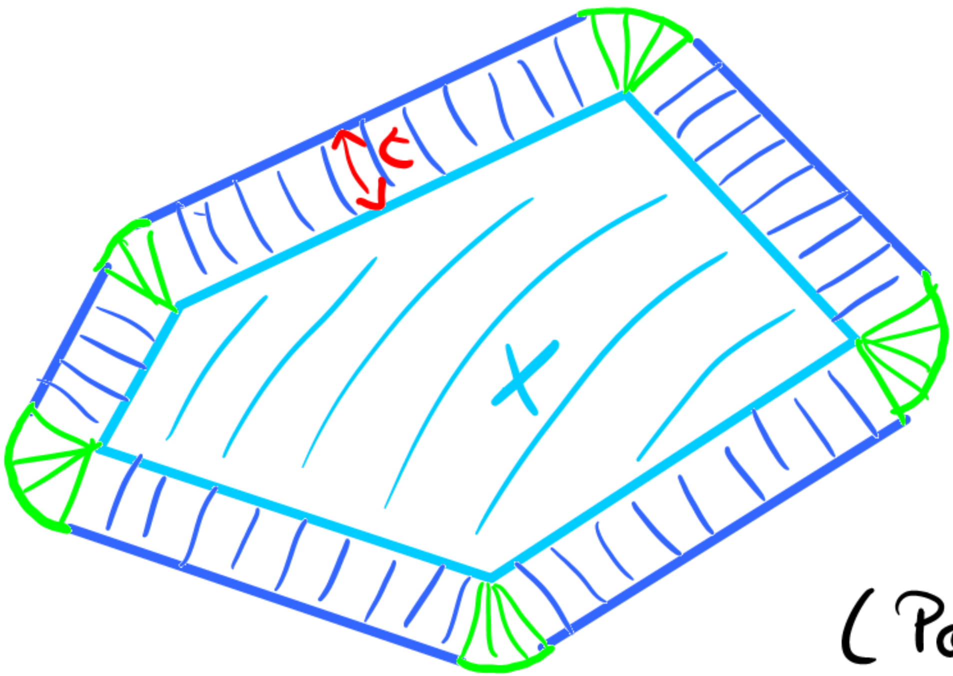
$$\mathrm{Vol}(X^t) = \sum_{i=0}^{d} \omega_i t^i V_{d-i}(X)$$

volume of the unit ball in $\mathbb{R}^i$.

Tube formula.

Within $[0, \mathrm{reach}(X)]$, (Federer, 1959)

$t \longmapsto \mathrm{Vol}(X^t)$ is a polynomial.



$$\mathrm{Vol}(X^t) = \mathrm{Vol}(X) + \mathrm{length}(\partial X)\, t + 2\pi\, \chi(X)\, t^2$$

(Polyhedra: Steiner, 1842)

Tube formula.

Within $[0, \text{reach}(X)]$, (Federer, 1959)

$t \mapsto \text{Vol}(X^t)$ is a polynomial.

When X is a smooth hypersurface of $\mathbb{R}^d$,

$$V_i(X) = \int_X \Sigma_{d-i-1}(\kappa_1, \ldots, \kappa_{d-1})\, d\tau$$ (Weyl, 1939)

symmetric polynomial   principal curvatures

Additivity

The additive property allows for a definition of intrinsic volumes for some sets of reach 0.

$$\boxed{V_i(A \cup B) + V_i(A \cap B) = V_i(A) + V_i(B)}.$$

$$V_i(+) = V_i(-) + V_i(|) - V_i(\cdot)$$

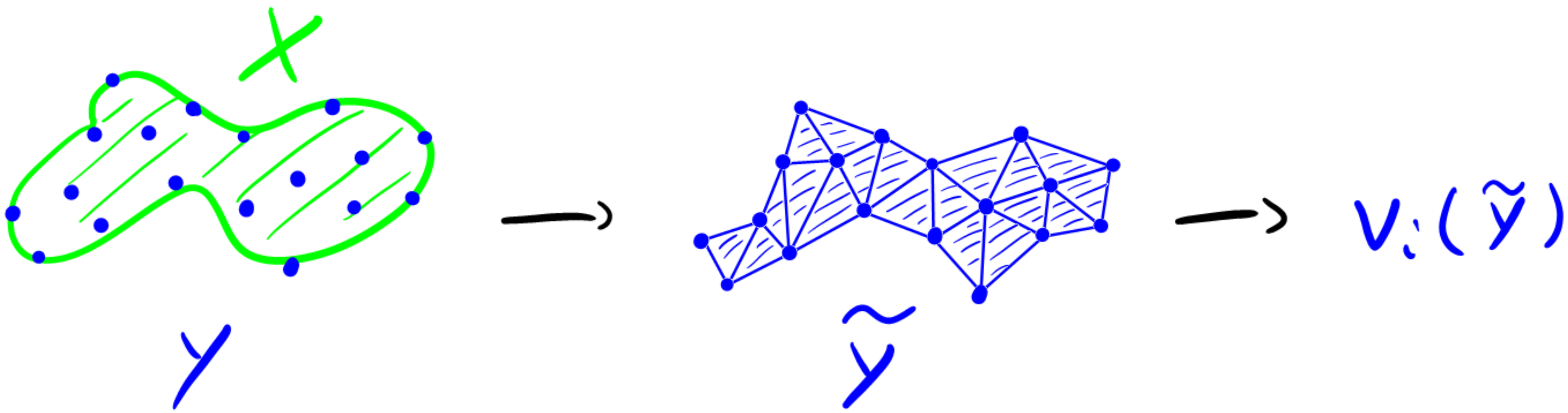
Inference on intrinsic volumes

Inference in the smooth case

When X is smooth,

intrinsic volumes can be recovered

by triangulating Y.



Such methods are supported by a vast literature.

Inference in the non-smooth case

When X is not smooth,

classical methods fail to reconstruct X.

Inference in the non-smooth case

When X is not smooth, classical methods fail to reconstruct X.

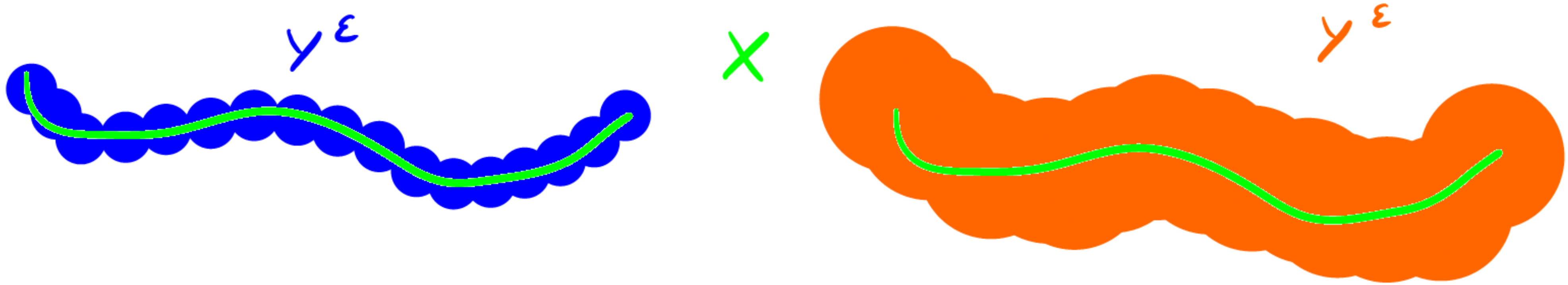
The only known method to reconstruct X consists in using offsets Y^ε. This recovers the homotopy type of X. (Chazal et al., 2007).

Inference in the non-smooth case

When X is not smooth,
classical methods fail to reconstruct X...

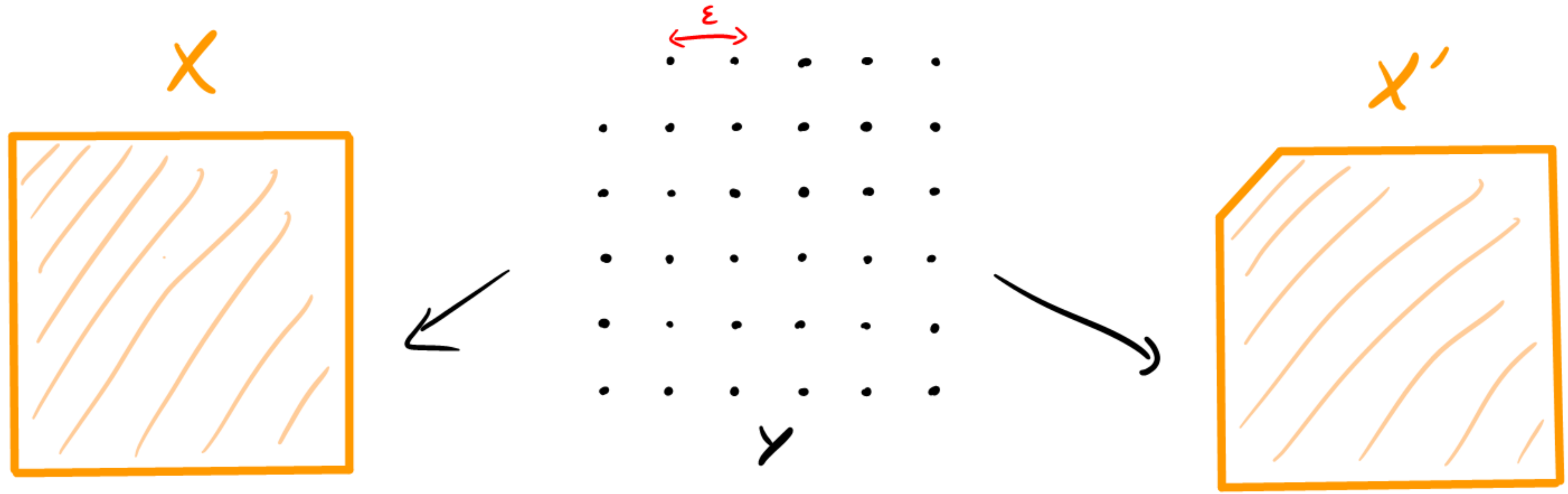
$V_i(Y^\varepsilon)$ fail to properly converge to $V_i(X)$ as either

- Y^ε is too noisy (ε is too small)
- Y^ε is smoother but far from X (ε is too large)



Inference in the non-smooth case

In the non-smooth case, the general rate of convergence with respect to the Hausdorff distance cannot be better than linear.



$d_H(X, Y), d_H(X', Y) = O(\varepsilon)$ but $\varepsilon = O(V_1(X) - V_1(X'))$

How do we deal with non-smooth sets?

Principal kinematic formula.

A special case of the known
Principal kinematic formula
yields

$$\int_{\mathbb{R}^d} \chi(X \cap B(x,t))\,dx = \sum_{i=0}^{d} \omega_i t^i V_{d-i}(X)$$

Holds for a large variety of sets.

Principal kinematic formula.

A special case of the known **Principal kinematic formula** yields

$$\int_{\mathbb{R}^d} \chi(X \cap B(x,t))\,dx = \sum_{i=0}^{d} \omega_i t^i V_{d-i}(X)$$

$\mathbb{1}_{X^t}(x)$ when $0 \leq t < \mathrm{reach}(X)$

Principal kinematic formula.

Idea: integrating the formula over t yields

$$\int_{\mathbb{R}^d} \int_0^R \chi\left(X \cap B(x,t)\right) dt\, dx$$

Principal kinematic formula.

Idea: integrating the formula over t yields

$$\boxed{\int_{\mathbb{R}^d}\int_0^R \chi(X \cap B(x,t))\,dt\,dx}$$

- $\forall x \in \mathbb{R}^d$, $(X \cap B(x,t))_{t\in\mathbb{R}}$ is a filtration
- The Euler characteristic is a topological quantity.

Principal kinematic formula.

Integrating the formula over t yields

$$\int_{\mathbb{R}^d}\int_0^R \chi\left(X \cap B(x,t)\right) dt\, dx$$

- $\forall x \in \mathbb{R}^d$, $\left(X \cap B(x,t)\right)_{t \in \mathbb{R}}$ is a filtration
- The Euler characteristic is a topological quantity.

Inviting us to use persistent homology.

Basics in homology and persistent homology.

Homology

- Let $i \in \mathbb{N}$ and $\mathbb{K}$ be a field.

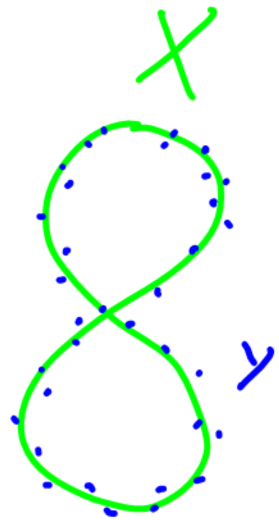
$H_i(X, \mathbb{K})$ is a vector space over $\mathbb{K}$.

Homology

• Let $i \in \mathbb{N}$ and $\mathbb{K}$ be a field.

$H_i(X, \mathbb{K})$ is a vector space over $\mathbb{K}$.

• $\dim H_i(X, \mathbb{K})$ is interpreted as the number of i-dimensional features (or voids) of X.



$\dim H_0(X)$ = number of connected components

$\dim H_1(X)$ = number of independent loops

Homology

- Let $i \in \mathbb{N}$ and $\mathbb{K}$ be a field.

$\underline{H_i(X, \mathbb{K})}$ is a vector space over $\mathbb{K}$.

- $\dim H_i(X, \mathbb{K})$ is interpreted as the number of <u>i-dimensional features (or voids)</u> of X.

- When the sum is well-defined,

$$\boxed{\chi(X) := \sum_{i=0}^{d} (-1)^i \dim H_i(X, \mathbb{K})}$$

Persistent homology : definition

A persistence module is a collection of vector spaces and linear maps.

$r \leq s \leq t$

$$\begin{tikzcd} \cdots \arrow[r, dashed] & M_r \arrow[r] \arrow[rr, bend left] & M_s \arrow[r] & M_t \arrow[r, dashed] & \cdots \end{tikzcd}$$

Persistent homology : definition

A persistence module is a collection of vector spaces and linear maps.

$$r \leq s \leq t$$

$$\begin{tikzcd} \cdots \arrow[r, dashed] & M_r \arrow[r] \arrow[rr, bend left] & M_s \arrow[r] & M_t \arrow[r, dashed] & \cdots \end{tikzcd}$$

<u>Example:</u> If $(X_t)_{t\in\mathbb{R}}$ is a filtration,

$$\cdots \dashrightarrow H_i(X_r) \longrightarrow H_i(X_s) \longrightarrow H_i(X_t) \dashrightarrow \cdots$$

We speak of persistent homology modules.

Persistent homology : decomposition

Under mild regularity conditions,
a persistence module can be decomposed
as a sum of interval modules $\mathbb{1}_I$

$$\cdots \dashrightarrow 0 \xrightarrow{0} \underbrace{\mathbb{K} \xrightarrow{\mathrm{id}} \mathbb{K}}_{I} \xrightarrow{0} 0 \dashrightarrow \cdots$$

Persistent homology : decomposition

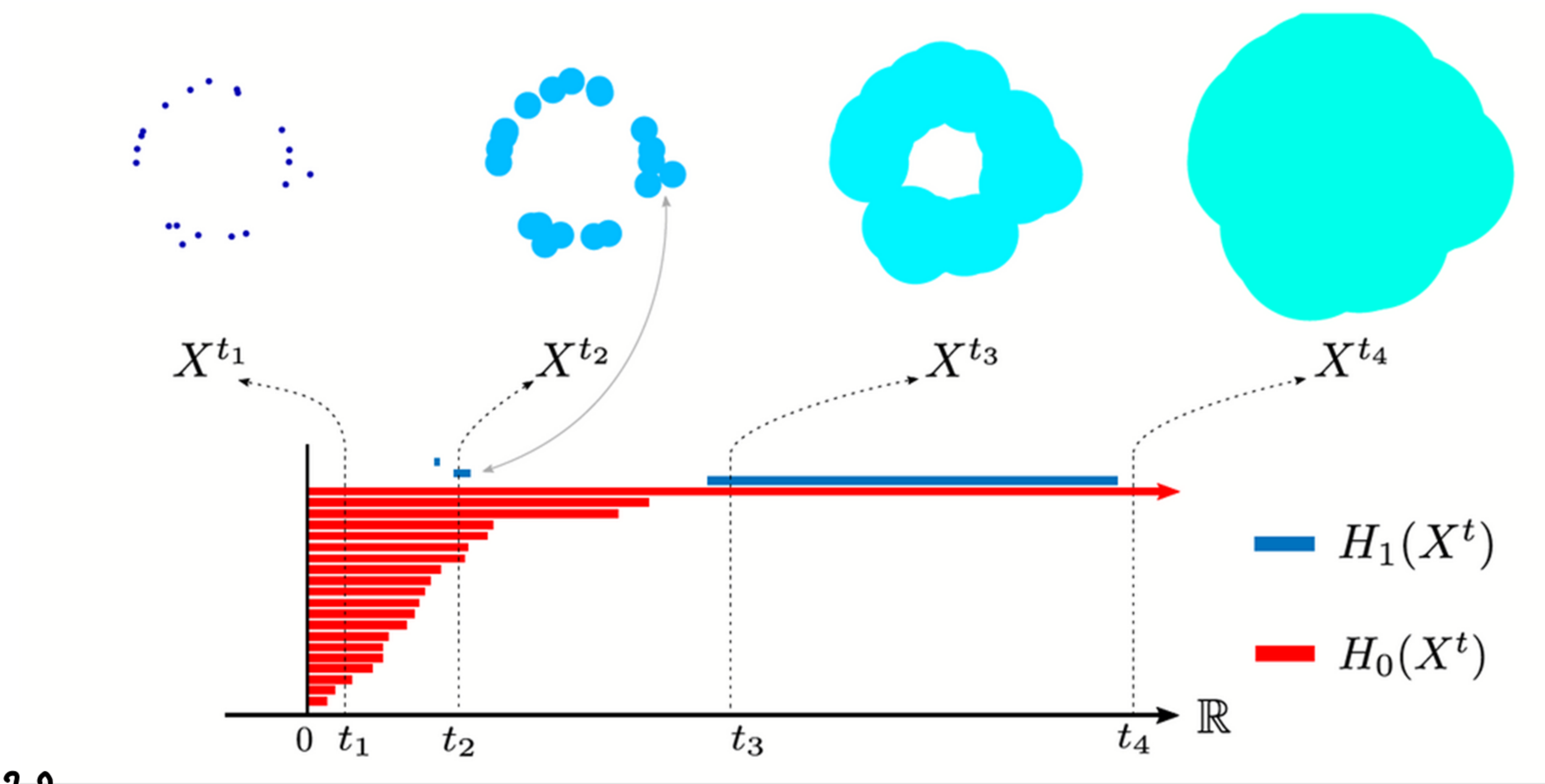
Under mild regularity conditions,
a persistence module can be decomposed
as a sum of interval modules $\mathbb{1}_I$

$$\cdots \to 0 \xrightarrow{0} \underbrace{\mathbb{K} \xrightarrow{id} \mathbb{K}}_{I} \xrightarrow{0} 0 \to \cdots$$

For persistent homology modules, each
interval's bounds correspond to the birth and death
of topological features.

Persistent homology : decomposition

Example: the growing offset filtration



Persistent homology: stability

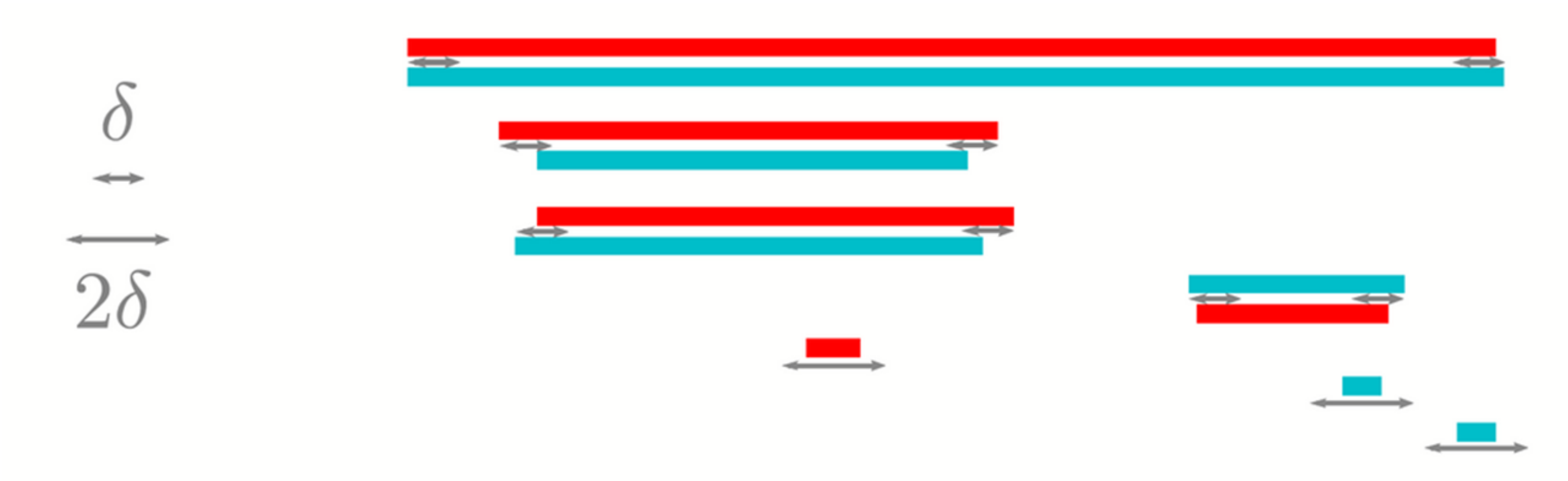
Persistence modules form a pseudo-metric space when equipped with the interleaving distance.

$d_I(M,N) :=$ infimum of δ such that

$$\begin{tikzcd}
\cdots \arrow[r, dashed] & M_t \arrow[r] \arrow[dr] & M_{t+\delta} \arrow[r] \arrow[dr] & M_{t+2\delta} \arrow[r, dashed] & \cdots \\
\cdots \arrow[r, dashed] & N_t \arrow[r] \arrow[ur] & N_{t+\delta} \arrow[r] \arrow[ur] & N_{t+2\delta} \arrow[r, dashed] & \cdots
\end{tikzcd}$$

Persistent homology: stability

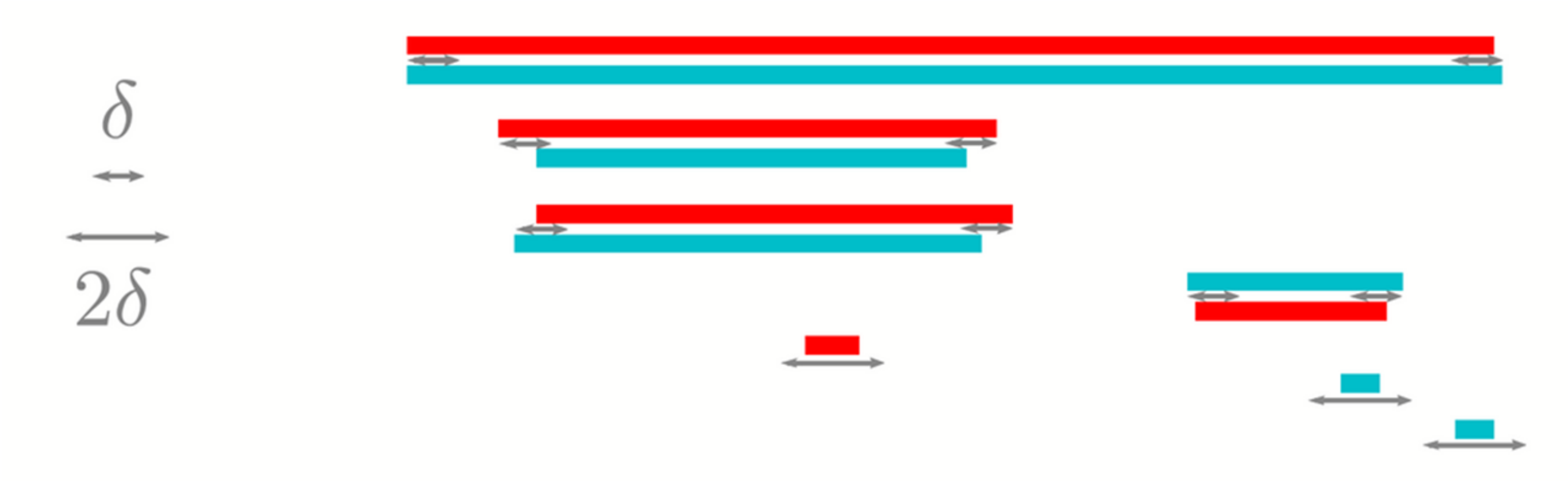
Persistence diagrams are equipped with the bottleneck distance d_B defined as the infimum of δ-matchings.



Partial bijection moving bounds by less than δ.

Persistent homology: stability

Persistence diagrams are equipped with the bottleneck distance d_B defined as the infimum of δ-matchings.



Isometry: $d_I(M,N) = d_B(dgm(M), dgm(N))$

Persistent homology: stability

Given $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ and $Z \subset \mathbb{R}^d$,
we let $\mathrm{dgm}(f_{|Z})$ be the persistence homology diagram associated with the filtration

$$\left(f^{-1}(-\infty, t]\right)_{t \in \mathbb{R}}$$

Persistent homology: stability

Given $f: \mathbb{R}^d \longrightarrow \mathbb{R}$ and $Z \subset \mathbb{R}^d$,
we let $\mathrm{dgm}(f_{|Z})$ be the persistence homology diagram associated with the filtration

$$\left(f^{-1}(-\infty, t]\right)_{t \in \mathbb{R}}$$

$\chi(\mathrm{dgm}_t(f_{|Z}))$ is the alternating sum of the number of intervals of $\mathrm{dgm}(f_{|Z})$ containing t

→ The Euler characteristic is defined for diagrams

Persistence

and

the kinematic formula.

Persistence and the kinematic formula.

The previous equation was:

$$\int_{\mathbb{R}^d} \chi(X \cap B(x,t))\,dx = \underbrace{\sum_{i=0}^{d} \omega_i t^i V_{d-i}(X)}_{Q_X(t)}$$

(Steiner Polynomial)

Persistence and the kinematic formula.

Let $d_x : z \longmapsto \|z - x\|$. The previous formula can be restated as

$$\int_{\mathbb{R}^d} \underbrace{\chi(\mathrm{dgm}_t(d_x|_X))dx}_{} = \underbrace{\sum_{i=0}^{d} \omega_i t^i V_{d-i}(X)}_{}$$

alternating sum of the number of intervals containing t.

$Q_X(t)$

(Steiner Polynomial)

Persistence and the kinematic formula.

$$\int_{\mathbb{R}^d} \chi\left(\mathrm{dgm}_t\left(d_{x|X}\right)\right)dx = \sum_{i=0}^{d} \omega_i t^i V_{d-i}(X)$$

Idea: approximate $\mathrm{dgm}(d_{x|X})$

Issue: $\mathrm{dgm}(d_{x|X})$, $\mathrm{dgm}(d_{x|Y})$, $\mathrm{dgm}(d_{x|Y^\varepsilon})$ can be very different.

Persistence and the kinematic formula.

<u>Idea</u>: Use two offsets of Y
to approximate $\mathrm{dgm}(d_x|_{X^{2\varepsilon}})$.

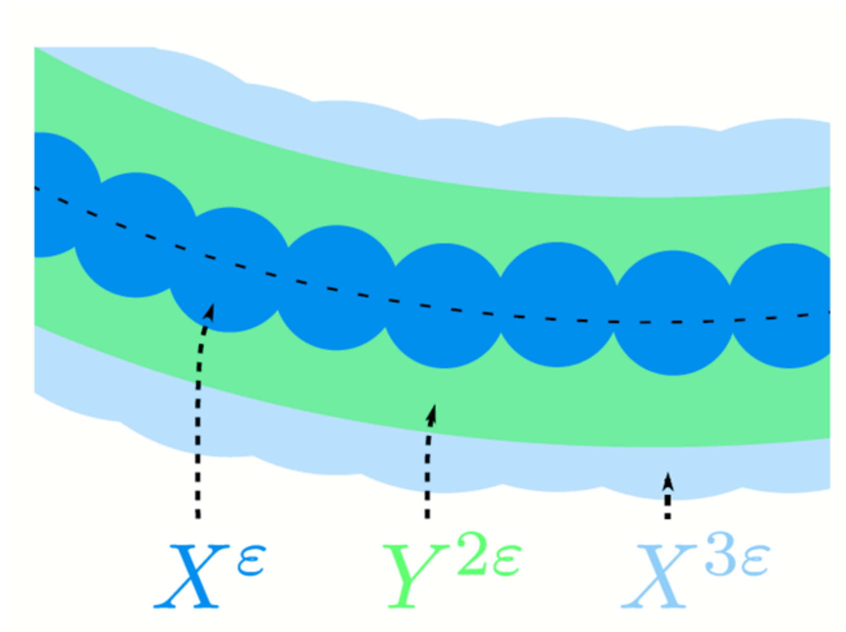


Image persistence

Fix $x \in \mathbb{R}^d$. Let $Z_t = Z \cap B(x,t)$ any set

$\forall \varepsilon > 0$, inclusions yield

a commutative diagram

$$\begin{tikzcd}
{} \arrow[r, dashed] & H_i(Y_t^{3\varepsilon}) \arrow[r] & H_i(Y_s^{3\varepsilon}) \arrow[r, dashed] & {} \\
{} \arrow[r, dashed] & H_i(Y_t^{\varepsilon}) \arrow[r] \arrow[u, "\iota_*"'] & H_i(Y_s^{\varepsilon}) \arrow[r, dashed] \arrow[u, "\iota_*"'] & {}
\end{tikzcd}$$

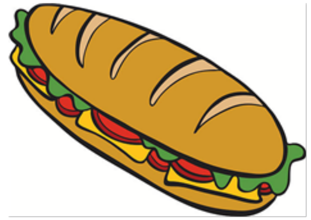
Image persistence

$$\begin{tikzcd}
\cdots \arrow[r, dashed] & H_i(Y_t^{3\varepsilon}) \arrow[r] & H_i(Y_s^{3\varepsilon}) \arrow[r, dashed] & \cdots \\
\cdots \arrow[r, dashed] & H_i(Y_t^{\varepsilon}) \arrow[r] \arrow[u, "\iota_*"'] & H_i(Y_s^{\varepsilon}) \arrow[r, dashed] \arrow[u, "\iota_*"'] & \cdots
\end{tikzcd}$$

$\Rightarrow \left(\iota_*(H_i(Y_t^{\varepsilon})) \right)_{t \in \mathbb{R}}$ is a **persistence module**.

Its diagram is denoted by $\mathrm{dgm}(d_x, Y^{\varepsilon}, Y^{3\varepsilon})$.

Image persistence



When $d_H(X, Y) \leq \varepsilon$, $\boxed{Y^{\varepsilon} \subset X^{2\varepsilon} \subset Y^{3\varepsilon}}$

$$
\begin{tikzcd}
\arrow[r, dashed] & H_i(Y^{3\varepsilon}_t) \arrow[r] & H_i(Y^{3\varepsilon}_\Delta) \arrow[r, dashed] & {} \\
\arrow[r, dashed] & H_i(X^{2\varepsilon}_t) \arrow[r] \arrow[u] & H_i(X^{2\varepsilon}_\Delta) \arrow[r, dashed] \arrow[u] & {} \\
\arrow[r, dashed] & H_i(Y^{\varepsilon}_t) \arrow[r] \arrow[u] \arrow[uu, bend left] & H_i(Y^{\varepsilon}_\Delta) \arrow[r, dashed] \arrow[u] \arrow[uu, bend right] & {}
\end{tikzcd}
$$

Image persistence

Theorem (Consequence of Bauer, 2013)

$\mathrm{dgm}(d_x, Y^{\varepsilon}, Y^{3\varepsilon})$ has fewer and smaller bars than $\mathrm{dgm}(d_{x|X^{2\varepsilon}})$

"Image persistent modules are simpler than the ones they sandwich"

Image persistence

Theorem (Consequence of Bauer, 2013)

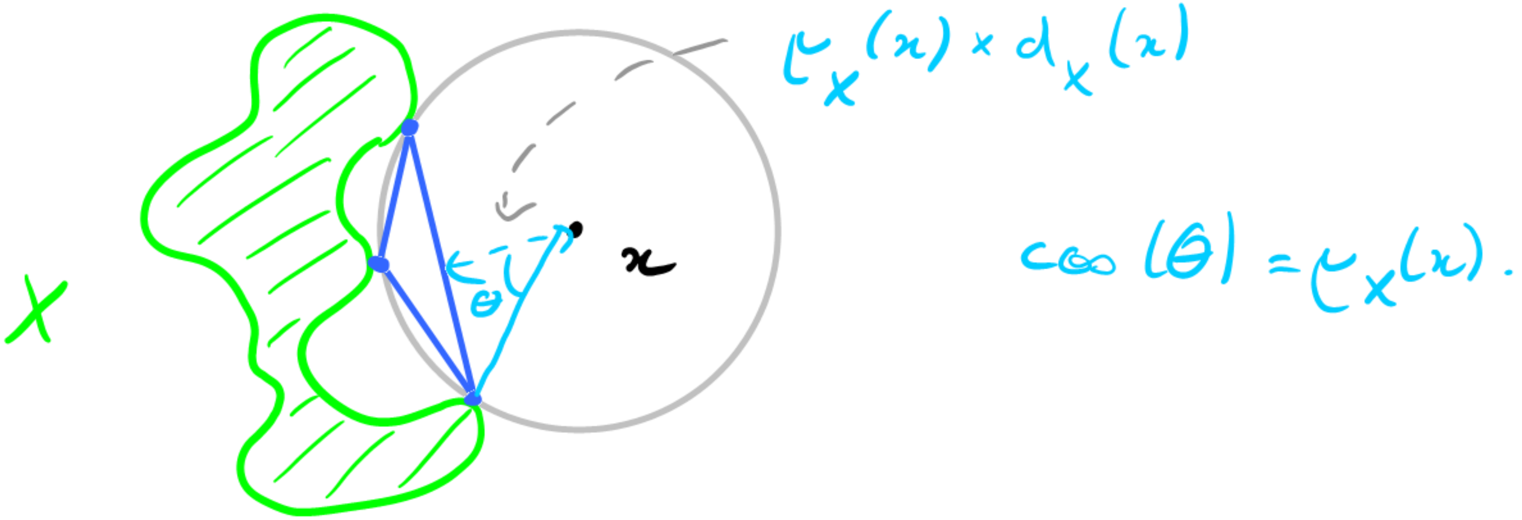
$\mathrm{dgm}(d_x, Y^{\varepsilon}, Y^{3\varepsilon})$ has fewer and smaller bars than $\mathrm{dgm}(d_{x|X^{2\varepsilon}})$

This can be seen as filtering the noise of individual offsets.

Regularity condition: μ-reach.

We need a mild regularity assumption.

For any $x \notin X$, define $\boxed{\mu_X(x)} \in [0,1]$ to be the distance of x to the convex hull of its closest points in X, normalized by $d_X(x)$.



Regularity condition: μ-reach.

We need a mild regularity assumption.

$$d_H(X, Y) \leq \varepsilon \leq \frac{1}{4}\,\mathrm{reach}_\mu(X)$$

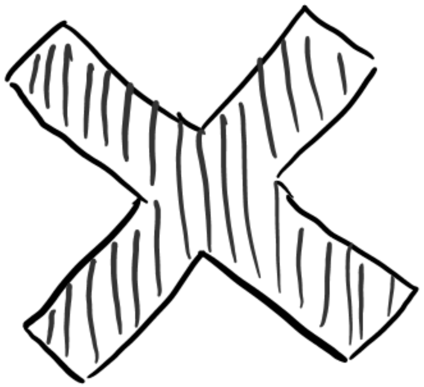
depending on a parameter $\mu \in [0,1]$.

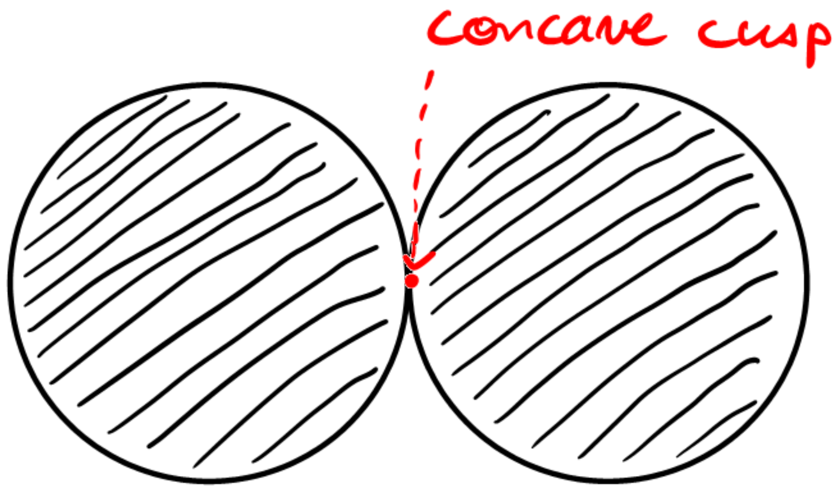
$$\mathrm{reach}_\mu(X) :=$$

$$\sup \left\{ t \in \mathbb{R},\ d_X(x) < t \implies \mu_X(x) \geq \mu \right\}$$

Regularity condition: μ-reach.

We need a mild regularity assumption.

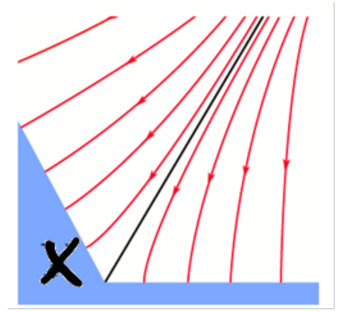




Positive μ-reach when $\mu \leq \frac{1}{\sqrt{2}}$

μ-reach zero for every μ in $(0,1]$.

Image stability



When $4\varepsilon \le \mathrm{reach}_\tau(X)$, there exist flows parametrized by the arc-length making d_X decrease at speed almost τ, yielding

$$\begin{tikzcd}
\cdots \arrow[r, dashed] & X_t^{4\varepsilon} \arrow[r] \arrow[dr, "\text{flow}"] & X^{4\varepsilon}_{t+\frac{2\varepsilon^+}{\tau}} \arrow[r, dashed] & \cdots \\
\cdots \arrow[r, dashed] & X_t^{2\varepsilon} \arrow[u] \arrow[r] \arrow[dr, "\text{flow}"] & X^{2\varepsilon}_{t+\frac{2\varepsilon^+}{\tau}} \arrow[u] \arrow[r, dashed] & \cdots \\
\cdots \arrow[r, dashed] & X_t \arrow[u] \arrow[r] & X_{t+\frac{2\varepsilon^+}{\tau}} \arrow[u] \arrow[r, dashed] & \cdots
\end{tikzcd}$$

Image stability

$$\left(c = \frac{2\varepsilon^+}{\tau}\right)$$

Applying the homology functor yields

$$\begin{tikzcd}
\cdots \arrow[r, dashed] & H_*(X^{4\varepsilon}_t) \arrow[r] \arrow[ddr] & H_*(X^{4\varepsilon}_{t+c}) \arrow[r] \arrow[ddr] & H_*(X^{4\varepsilon}_{t+2c}) \arrow[r, dashed] & \cdots \\
\cdots \arrow[r, dashed] & H_*(Y^{3\varepsilon}_t) \arrow[r] \arrow[u] & H_*(Y^{3\varepsilon}_{t+c}) \arrow[r] \arrow[u] & H_*(Y^{3\varepsilon}_{t+2c}) \arrow[r, dashed] \arrow[u] & \cdots \\
\cdots \arrow[r, dashed] & H_*(X^{2\varepsilon}_t) \arrow[r] \arrow[u] \arrow[ddr] & H_*(X^{2\varepsilon}_{t+c}) \arrow[r] \arrow[u] \arrow[ddr] & H_*(X^{2\varepsilon}_{t+2c}) \arrow[r, dashed] \arrow[u] & \cdots \\
\cdots \arrow[r, dashed] & H_*(Y^{\varepsilon}_t) \arrow[r] \arrow[u] & H_*(Y^{\varepsilon}_{t+c}) \arrow[r] \arrow[u] & H_*(Y^{\varepsilon}_{t+2c}) \arrow[r, dashed] \arrow[u] & \cdots \\
\cdots \arrow[r, dashed] & H_*(X_t) \arrow[r] \arrow[u] & H_*(X_{t+c}) \arrow[r] \arrow[u] & H_*(X_{t+2c}) \arrow[r, dashed] \arrow[u] & \cdots
\end{tikzcd}$$

Image stability

$$\left(c = \frac{2\varepsilon^+}{\tau} \right)$$

$$\begin{tikzcd}
{} \arrow[r, dashed] & H_*(X_t^{4\varepsilon}) \arrow[r] & H_*(X_{t+c}^{4\varepsilon}) \arrow[r] \arrow[ddr] & H_*(X_{t+2c}^{4\varepsilon}) \arrow[r, dashed] & {} \\
{} \arrow[r, dashed] & H_*(Y_t^{3\varepsilon}) \arrow[r] \arrow[dr, bend left] & H_*(Y_{t+c}^{3\varepsilon}) \arrow[u] \arrow[r] & H_*(Y_{t+2c}^{3\varepsilon}) \arrow[u] \arrow[r, dashed] & {} \\
{} \arrow[r, dashed] & H_*(X_t^{2\varepsilon}) \arrow[u] \arrow[r] \arrow[ddr] & H_*(X_{t+c}^{2\varepsilon}) \arrow[u] \arrow[r] \arrow[dr, bend right] & H_*(X_{t+2c}^{2\varepsilon}) \arrow[u] \arrow[r, dashed] & {} \\
{} \arrow[r, dashed] & H_*(Y_t^{\varepsilon}) \arrow[u] \arrow[r] & H_*(Y_{t+c}^{\varepsilon}) \arrow[u] \arrow[r] & H_*(Y_{t+2c}^{\varepsilon}) \arrow[u] \arrow[r, dashed] & {} \\
{} \arrow[r, dashed] & H_*(X_t) \arrow[u] \arrow[r] & H_*(X_{t+c}) \arrow[u] \arrow[r] & H_*(X_{t+2c}) \arrow[r, dashed] & {}
\end{tikzcd}$$

Image stability

Conclusion: [C, C-S]

Assume $d_H(X,Y) \leq \varepsilon \leq \frac{1}{4} \mathrm{reach}_C(X)$

$$d_B\left(\mathrm{dgm}\left(d_{x|X^{2\varepsilon}}\right), \mathrm{dgm}\left(d_x, Y^{\varepsilon}, Y^{3\varepsilon}\right)\right) \leq \frac{2\varepsilon}{C}$$

Back to intrinsic volumes

Back to intrinsic volumes.

Recall that we have

$$\int_{\mathbb{R}^d} \chi\left(\mathrm{dgm}_t\left(d_{x|X^{2\varepsilon}}\right)\right) dx = \underbrace{\sum_{i=0}^{d} \omega_i t^i V_{d-i}(X^{2\varepsilon})}_{Q_{X^{2\varepsilon}}(t)}$$

alternating sum of the number of intervals containing t.

<u>Idea</u>: replace $\mathrm{dgm}(d_{x|X^\varepsilon})$ by $\mathrm{dgm}(d_x, Y^\varepsilon, Y^{3\varepsilon})$

Back to intrinsic volumes.

We let Q_Y^{ε} be the persistent Steiner function

$$Q_Y^{\varepsilon}(t) := \int_{\mathbb{R}^d} \chi\left(\mathrm{dgm}\left(d_x, Y^{\varepsilon}, Y^{3\varepsilon}\right)\right) dx$$

Back to intrinsic volumes.

To compare $Q_{X^{2\varepsilon}}$ and Q_Y^{ε}, we use

χ-averaging Lemma:

$$\int_0^R \left| \chi(\mathrm{dgm}_t(d_{x|X^{2\varepsilon}})) - \chi(\mathrm{dgm}_t(d_x, Y^{\varepsilon}, Y^{3\varepsilon})) \right| dt \leq 2\, d_B \cdot N_0^R(X^{2\varepsilon}, x)$$

distance between diagrams

where $N_0^R(X^{2\varepsilon}, x)$ is the number of bars of $\mathrm{dgm}(d_{x|X^{2\varepsilon}})$ intersecting $[0,R]$.

Convergence bounds.

Since $d_B \leq \frac{2\varepsilon}{c}$, we have

Corollary:

$$\underbrace{\|Q_{X^{2\varepsilon}} - Q_{Y,\varepsilon}\|_{1,[0,R]} \leq 4\frac{\varepsilon}{c}\int_{\mathbb{R}^d} N_0^R(X^{2\varepsilon},x)\,dx}$$

$$\underbrace{\int_{\mathbb{R}^d} N_0^R(X^{2\varepsilon},x)\,dx}_{K_R(X^{2\varepsilon})}$$

Convergence bounds

Moreover, the same method yields

$$\|Q_{X^{2\varepsilon}} - Q_{X^{\delta}}\|_{1,[0,R]} \leq 4\frac{\varepsilon}{\tau}\int_{\mathbb{R}^d}\left(N_0^R(X^\delta, x) + N_0^R(X^{2\varepsilon}, x)\right)dx$$

Letting δ go to zero* yields

Inference bounds [C.]

$$\|Q_X - Q_{Y,\varepsilon}\|_{1,[0,R]} \leq 4\frac{\varepsilon}{\tau}\left(K_R(X) + K_R(X^{2\varepsilon})\right)$$

Convergence bounds

Recovering surrogate coefficients of $Q_{Y,\varepsilon}$ can be done in a linear, continuous way, yielding

$$|V_{i,\varepsilon}(Y) - V_i(X^{2\varepsilon})| = O\left(\frac{\varepsilon}{\tau} K(X^{2\varepsilon})\right)$$

surrogates

&

$$|V_{i,\varepsilon}(Y) - V_i(X)| = O\left(\frac{\varepsilon}{\tau}\left(K_R(X) + K_R(X^{2\varepsilon})\right)\right)$$

What can we say about $K_{\mathbb{R}}(X^{2\varepsilon})$?

What can we say about $K_R(X^{2\varepsilon})$?

Recall that $K_R(X^{2\varepsilon}) = \int_{\mathbb{R}^d} N_0^R(X^{2\varepsilon}, x)\,dx$

Morse theory.

$N_0^R(X^{2\varepsilon}, x)$ is the number of bars of $\operatorname{dgm}(d_x|_{X^{2\varepsilon}})$ intersecting $[0,R]$.

Morse theory.

$N_0^R(X^{2\varepsilon}, x)$ is the number of bars of $\text{dgm}(d_{x|X^{2\varepsilon}})$ intersecting $[0,R]$.

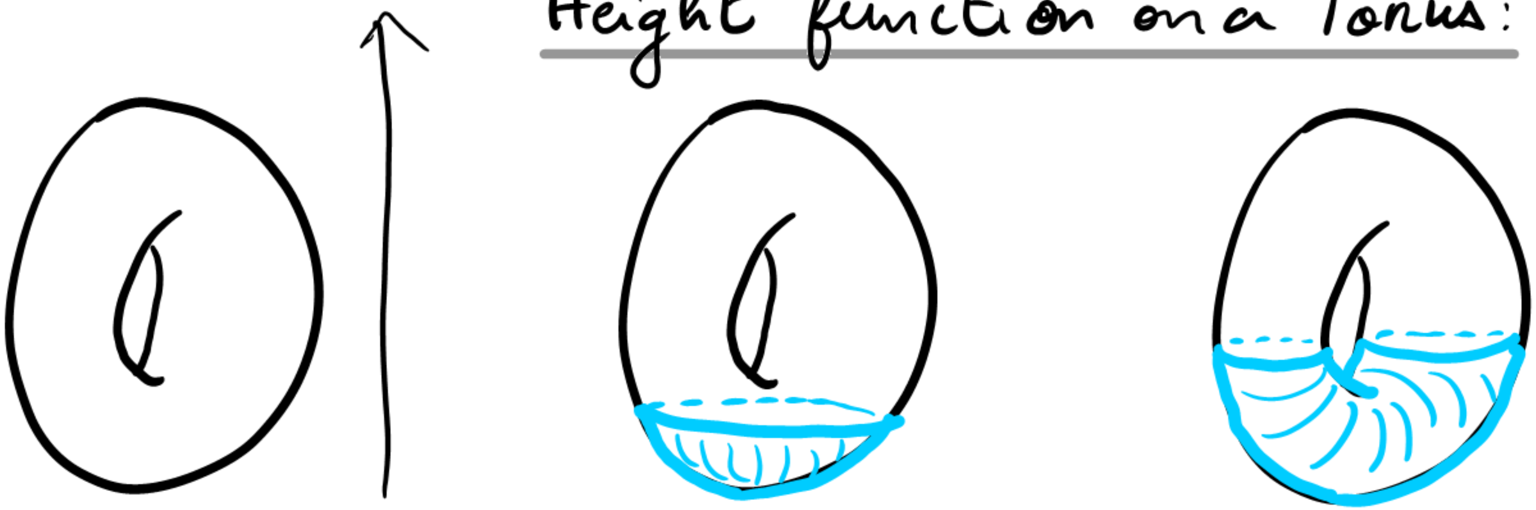
$\text{dgm}(d_{x|X^{2\varepsilon}})$ is the persistent diagram associated with the filtration $(X^{2\varepsilon} \cap B(x,t))_{t\in\mathbb{R}}$

That is, the sublevel set filtration $d_x^{-1}(-\infty, t]$ of $d_{x|X^{2\varepsilon}} : \begin{cases} X^{2\varepsilon} \longrightarrow \mathbb{R}^+ \\ y \longmapsto \|y - x\| \end{cases}$

Morse theory.

When Z is smooth and $f|_Z : Z \longrightarrow \mathbb{R}$ is a Morse function, the number of bars intersecting $[0,R]$ is bounded by the number of critical points of $f|_Z$

Height function on a Torus:



Morse theory.

When Z is smooth and $f_{|Z} : Z \longrightarrow \mathbb{R}$ is a Morse function, the number of bars intersecting $[0,R]$ is bounded by the number of critical points of $f_{|Z}$

$\Rightarrow$ if $d_{x|X^{2\varepsilon}}$ has a Morse function behavior, (even though $X^{2\varepsilon}$ is not smooth), it suffices to study the critical points of $d_{x|X^{2\varepsilon}}$.

Morse theory.

Inspired by Joseph Fu, we developped a theory of Morse functions restricted to offsets X^δ.

when δ is a regular value of d_X.

Morse theory.

Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a C^2 map.

Taking inspiration from Fu, we define critical points and Hessian of $f_{|X^\delta}$ depending on f and the curvatures of X^δ.

Morse theory.

Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a C^2 map.

taking inspiration from Fu, we define

critical points and Hessian of $f|_{X^\delta}$

depending on f and the curvatures of X^δ.

f is said to be Morse when

at every critical point, the Hessian

is non-degenerate.

Morse theory.

Theorem: [C.]

Let $f : \mathbb{R}^d \to \mathbb{R}$ C^2 be such that $f|_{X^\delta}$ is Morse.

Then the filtration $(f^{-1}(-\infty, t] \cap X^\delta)_{t \in \mathbb{R}}$ is such that

- Between critical values, the homotopy type stays constant.

Morse theory.

Theorem: [C.]

Let $f: \mathbb{R}^d \to \mathbb{R}$ C^2 be such that

$f|_{X^\delta}$ is Morse.

Then the filtration $\left(f^{-1}(-\infty, t] \cap X^\delta\right)_{t \in \mathbb{R}}$

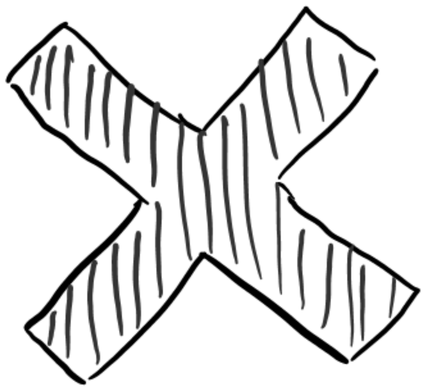
is such that

- Between critical values, the homotopy type stays constant.

- Around a critical point, a cell is added, corresponding to one event in the associated diagram.

Morse theory.

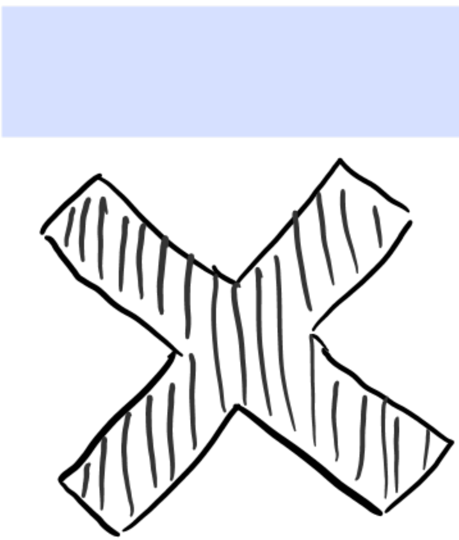
Example.



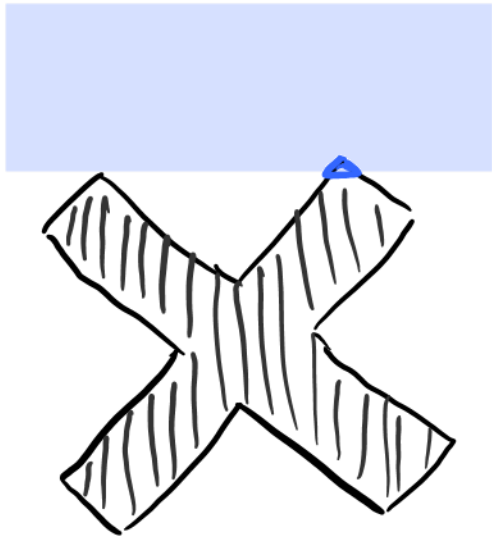
$\emptyset$

Morse theory.

Example.



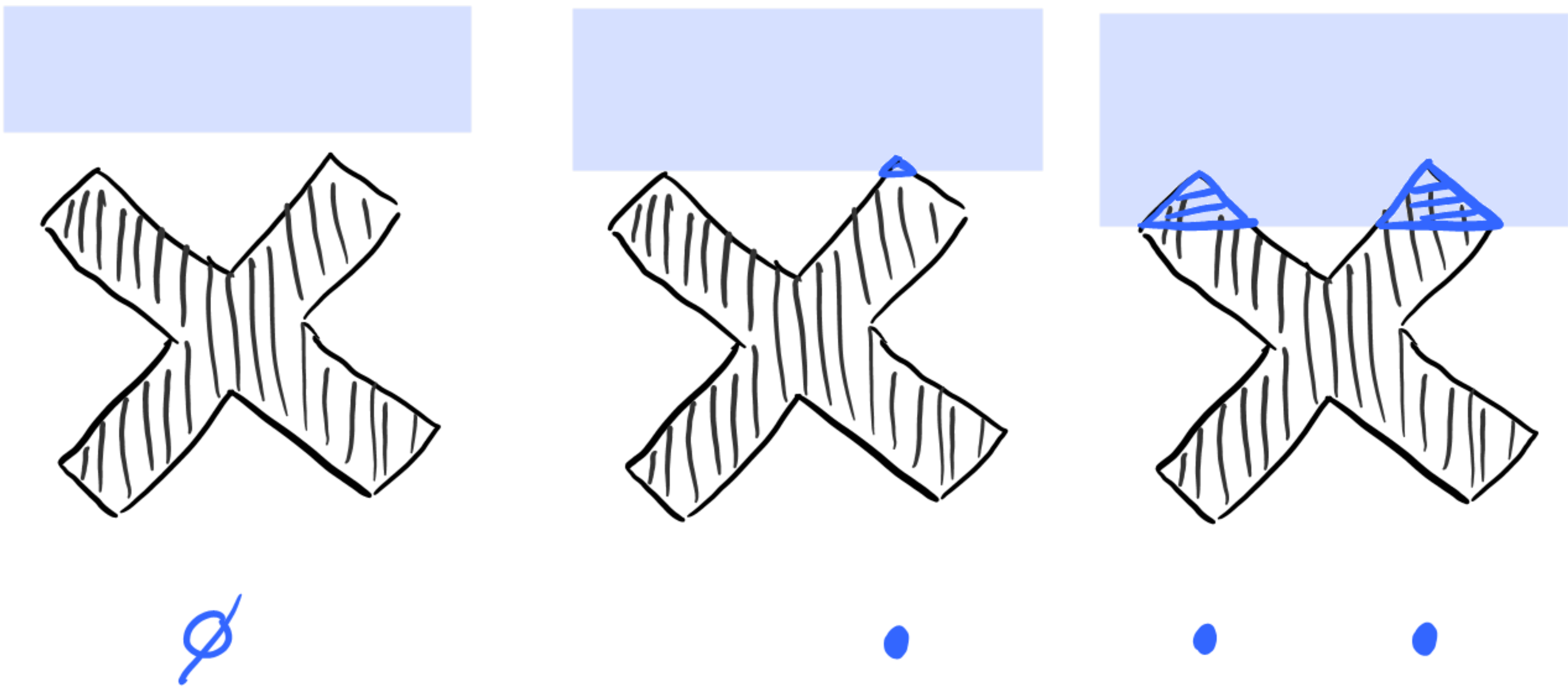
$\emptyset$



$\bullet$

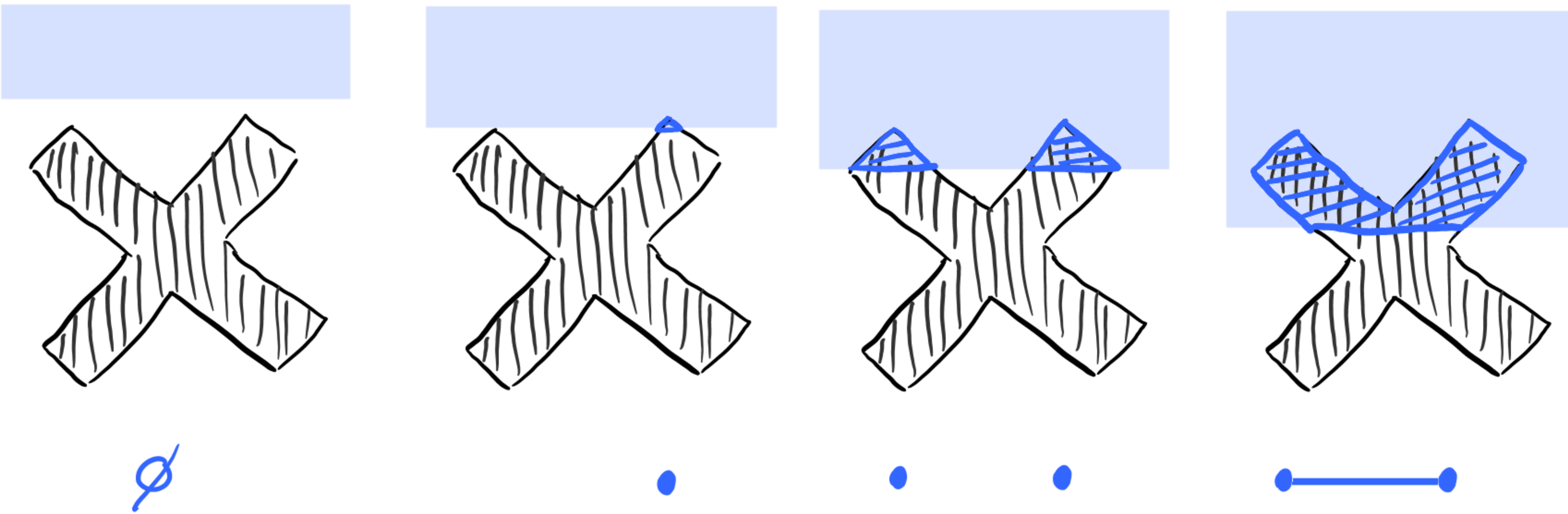
Morse theory.

Example.



Morse theory.

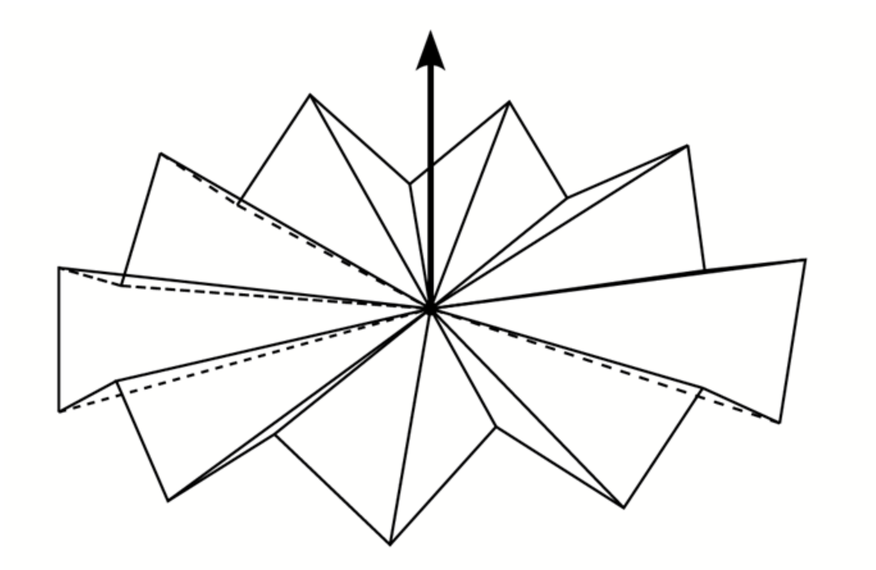
Example.

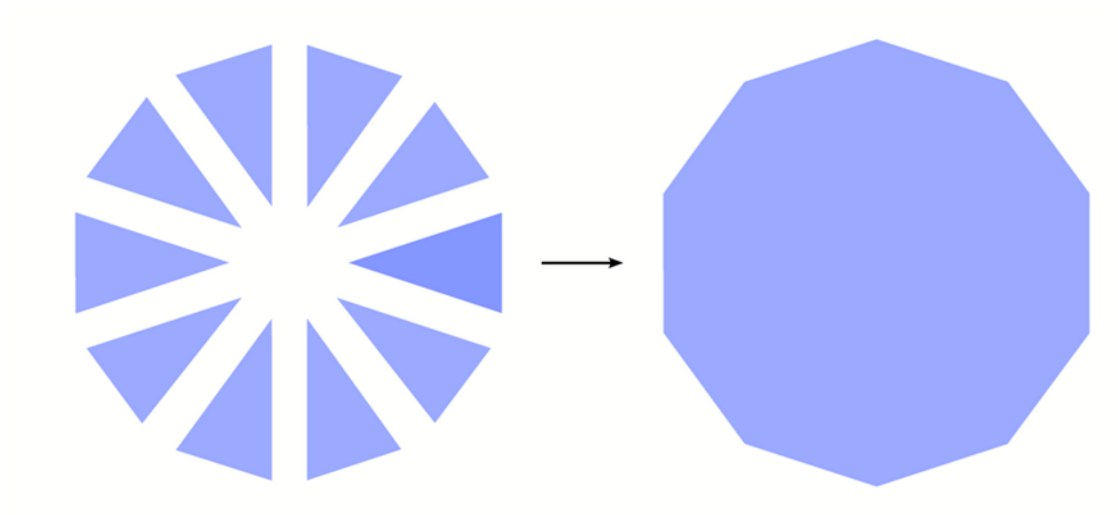


Persistence diagram:

Morse theory.

Counter example





One critical point, several changes.

Morse theory.

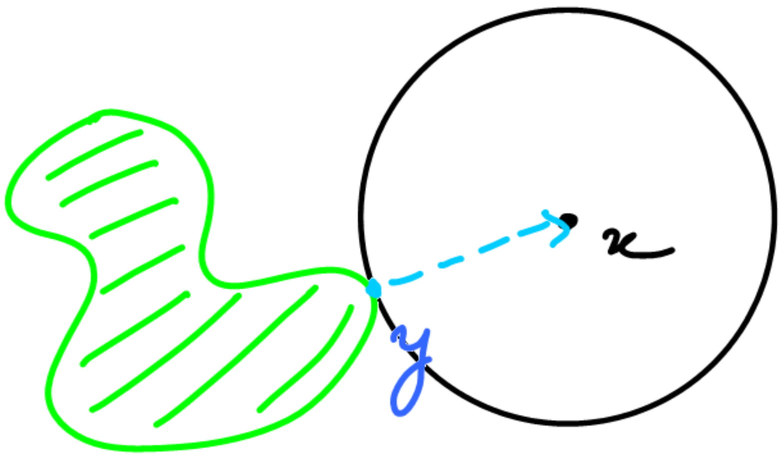
Proposition: [C.]

For almost every x in $\mathbb{R}^d$, $d_{x|X^\delta}$ is a Morse function.

Morse theory.

Proposition: [C.]

For almost every x in $\mathbb{R}^d$, $d_{x|X^\delta}$ is a Morse function.



Critical points of $d_{x|X}$ are related to the normals of X.

Bound on $K_R(X^\delta)$.

Bounding $N_0^R(X^\delta, x)$ by the number of critical points of $d_{x|X^\delta}$, and integrating this number yields

Function of the curvatures of X^δ.

$$\boxed{K_R(X^\delta) \leq \mathrm{Vol}(X^\delta) + \overbrace{M_R(X^\delta)}}$$

$$\Rightarrow \quad \boxed{K_R(X^\delta) = O(\mathrm{Vol}(X^\delta) + \underbrace{M(N_X)})}$$

Mass of the unit normal bundle ($\simeq$ total curvatures)

Side result: stability of intrinsic volumes

Let $X, Y \subset \mathbb{R}^d$.

Assume they are ε-homotopy equivalent, i.e

$$\exists\, f : X \longrightarrow Y, \quad g : Y \longrightarrow X$$

such that $f \circ g$, $g \circ f$ are homotopic to Id_Y, Id_X with homotopy trajectories bounded by ε.

Then $\boxed{|V_i(X) - V_i(Y)| = O(\varepsilon(K(X) + K(Y)))}$

Conclusion

- Using tools from persistent homology, geometric measure theory and non-smooth analysis, we were able to extend the noise-filtering property of persistence theory to the realm of geometry and obtain inference results on non-smooth sets converging at the optimal rate.

$$|V_{i,\varepsilon}(Y) - V_i(X)| = O\left(\frac{\varepsilon}{r}\left(K(X) + K(X^{2\varepsilon})\right)\right)$$

- Along the way, we obtained a result on Morse theory and a stability result for intrinsic volumes.

Open problems

- Intrinsic volumes are global curvatures.

  It is possible to recover the curvature measures of X from Y?

  Or its normal cycle?

- Mimicking what we did with $Y^{\varepsilon} \hookrightarrow Y^{3\varepsilon}$, can we exploit any inclusion $A \hookrightarrow B$?

Danke!

Grazie!

Thank you for your attention!

Merci!

Děkuji!

მადლობა!