Subdivide and Tile:
Triangulating spaces for understanding the world
Leiden, Nov. 2009

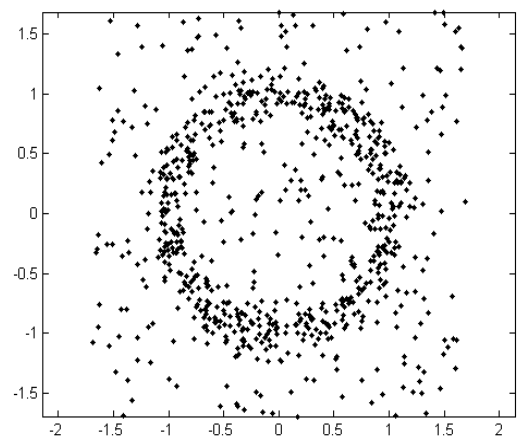
Geometric Inference

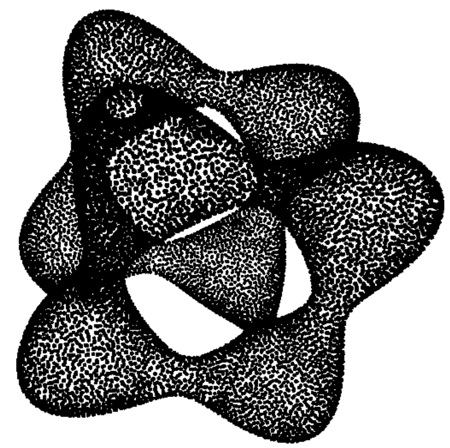
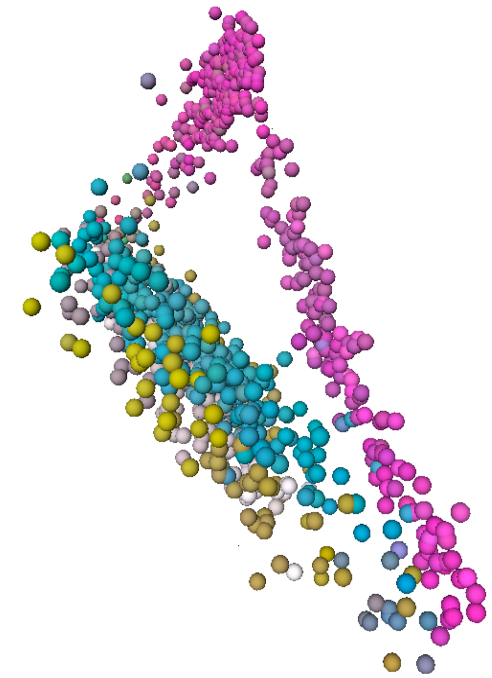
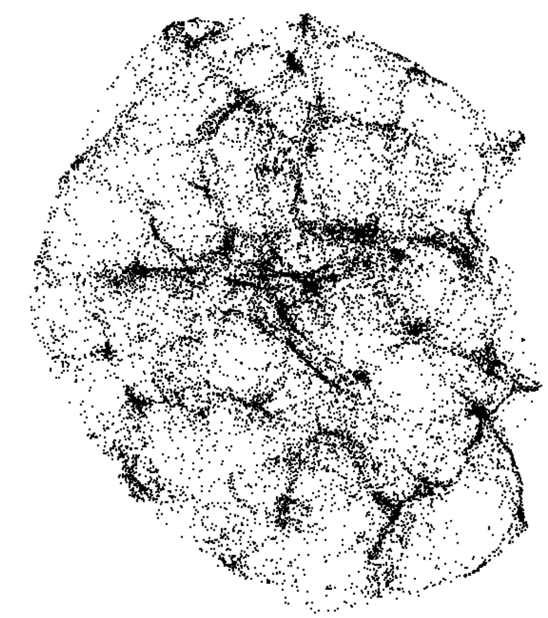
F. Chazal
Geometrica Group
INRIA Saclay

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
INRIA
centre de recherche
SACLAY - ÎLE-DE-FRANCE

GEOMETRICA

Introduction and motivations



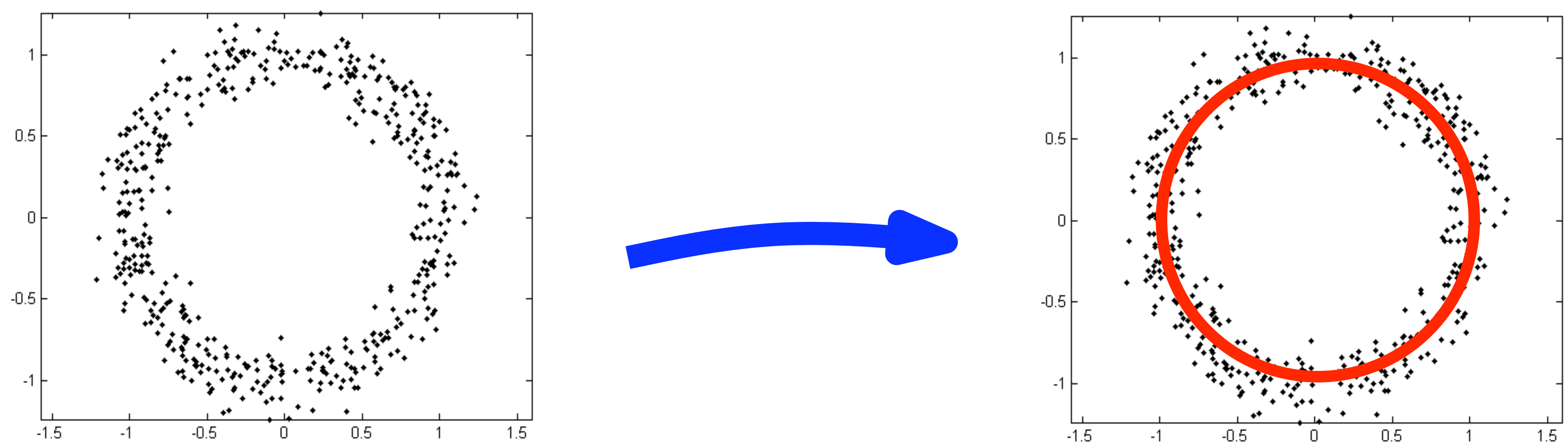




What can we say about the topology/geometry of spaces known only through a finite set of measurements?

What is the relevant topology/geometry of a point cloud data set?

Motivations: Reconstruction, Manifold Learning and NLDR, Clustering and Segmentation,...

Geometric Inference



Question: Given an approximation C of a geometric object K, is it possible to reliably estimate the topological and geometric properties of K, knowing only the approximation C?

Question *: Given a point cloud C (or some other more complicated set), is it possible to infer some robust topological or geometric information of C?

- The answer depends on:
  - the considered class of objects (no hope to get a positive answer in full generality),
  - a notion of distance between the objects (approximation).

Outline

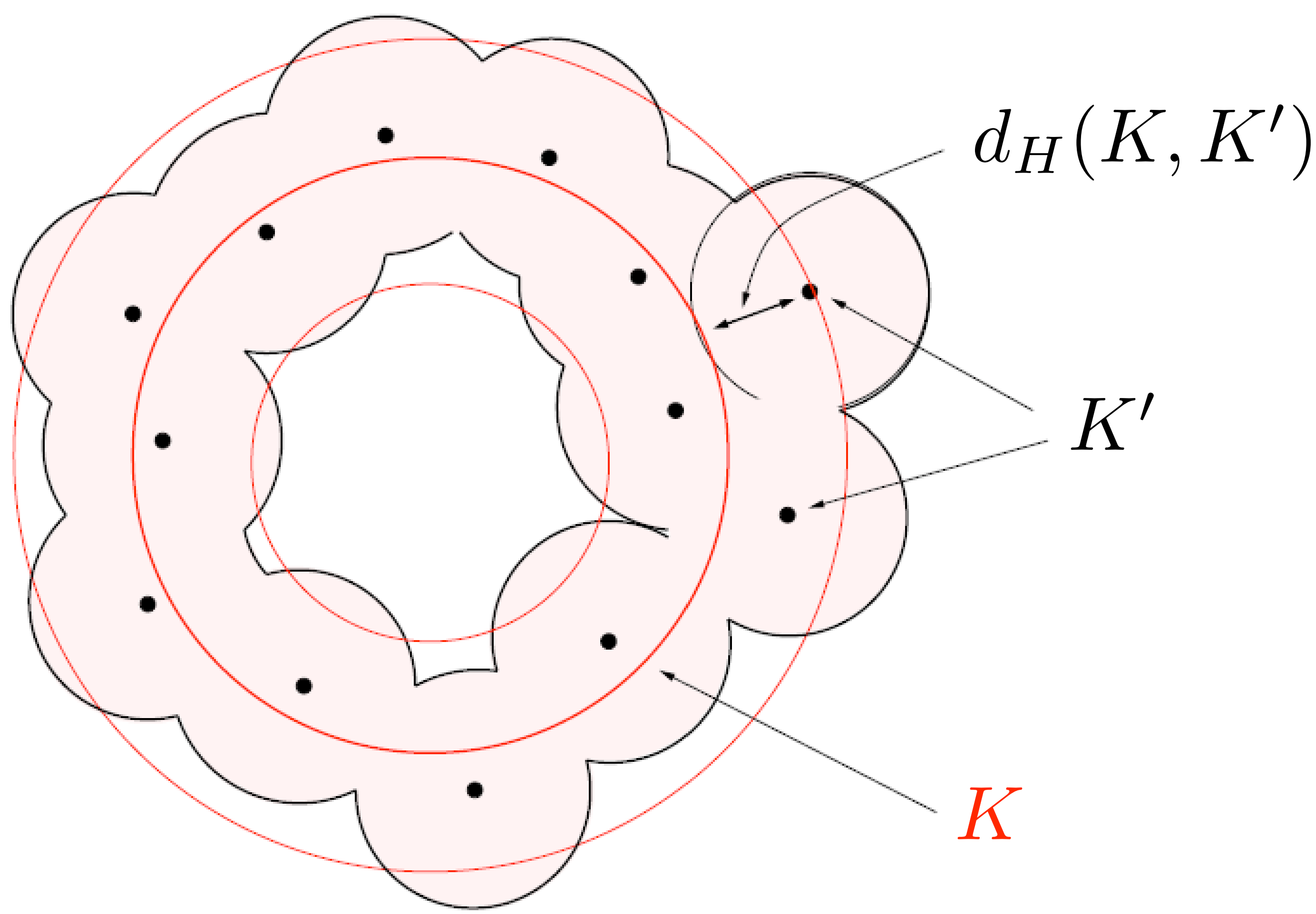
1. Distance functions for geometric inference
   - class of objects: (some) compact subsets of $\mathbb{R}^d$
   - approximation with respect to Hausdorff distance
2. (Practical) algorithms for topological inference
   - persistence-based algorithm
   - multiscale inference
3. Dealing with outliers: the measure point of view
   - class of objects: probability measures
   - approximation with respect to Wasserstein distance

Distance functions for geometric inference

Considered objects: compact subsets K of $\mathbb{R}^d$

Distance:
distance function to a compact $K \subset \mathbb{R}^d$: $d_K : x \to \inf_{p \in K} \|x - p\|$
Hausdorf distance between two compact sets:

$$d_H(K, K') = \sup_{x \in \mathbb{R}^d} |d_K(x) - d_{K'}(x)|$$



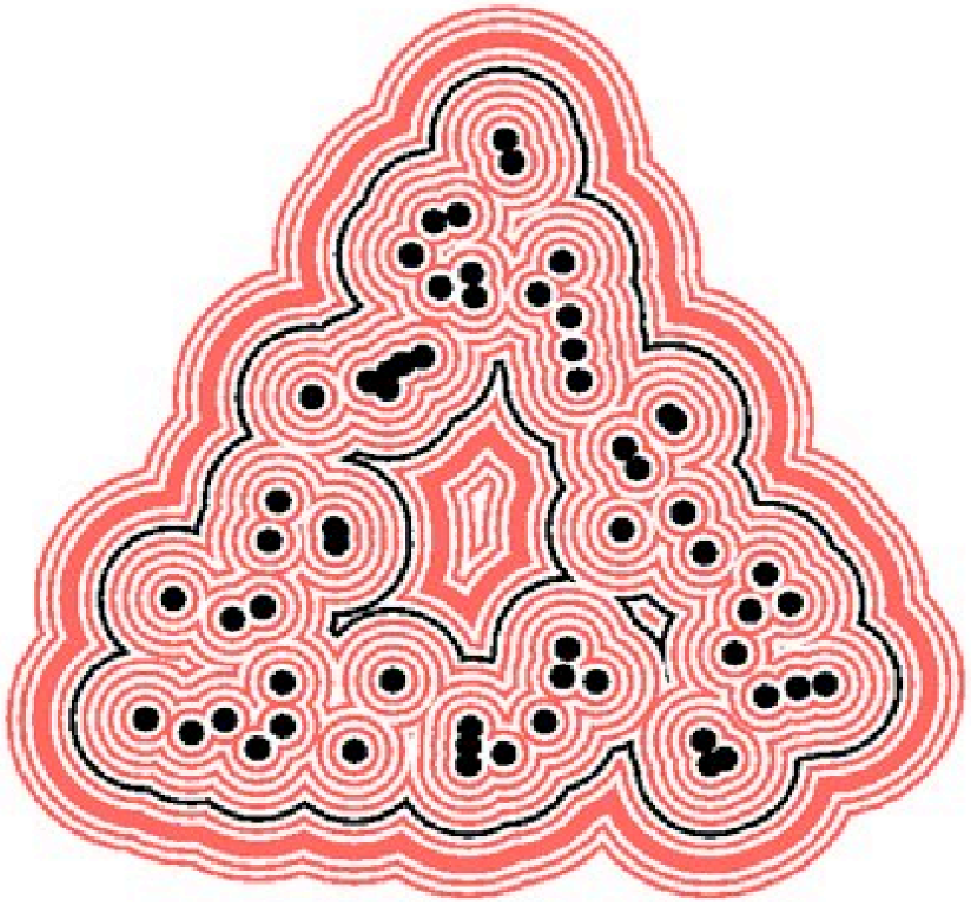
Distance functions for geometric inference

Considered objects: compact subsets K of $\mathbb{R}^d$

Distance:
distance function to a compact $K \subset \mathbb{R}^d$: $d_K : x \to \inf_{p \in K} \|x - p\|$
Hausdorf distance between two compact sets:

$$d_H(K, K') = \sup_{x \in \mathbb{R}^d} |d_K(x) - d_{K'}(x)|$$

- Replace K and C by d_K and d_C
- Compare the topology of the offsets $K^r = d_K^{-1}([0, r])$ and $C^r = d_C^{-1}([0, r])$

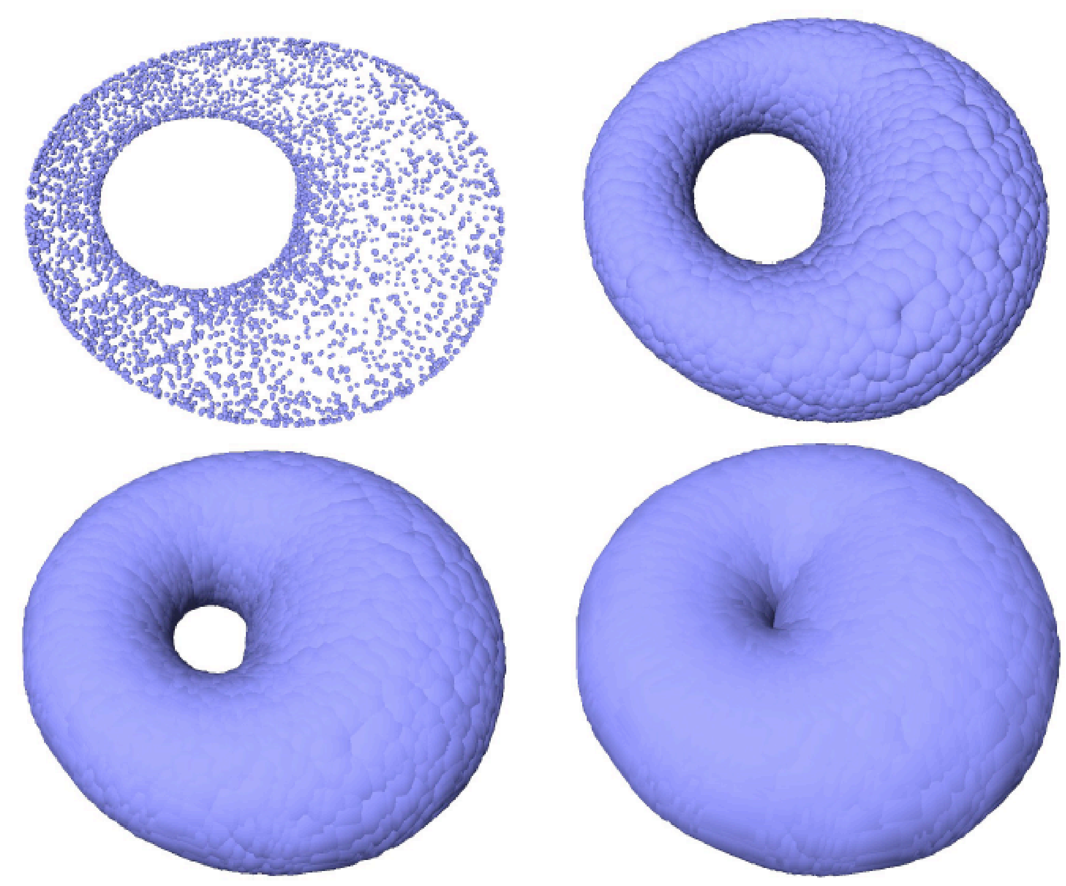


Distance functions for geometric inference

Considered objects: compact subsets K of $\mathbb{R}^d$

Distance:
distance function to a compact $K \subset \mathbb{R}^d$: $d_K : x \to \inf_{p \in K} \|x - p\|$
Hausdorf distance between two compact sets:

$$d_H(K, K') = \sup_{x \in \mathbb{R}^d} |d_K(x) - d_{K'}(x)|$$



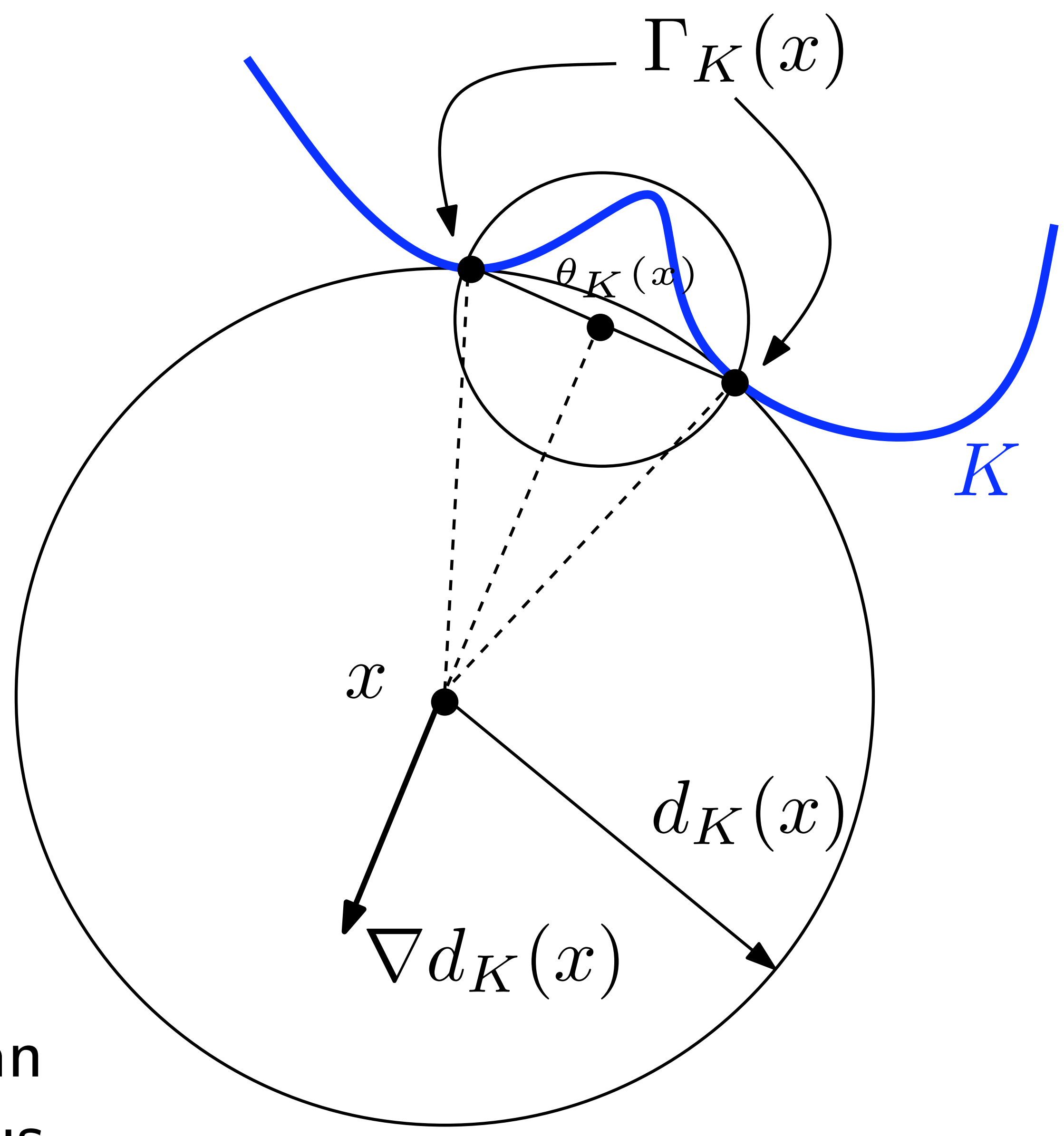
- Replace K and C by d_K and d_C
- Compare the topology of the offsets $K^r = d_K^{-1}([0, r])$ and $C^r = d_C^{-1}([0, r])$

The gradient of the distance function

- $\Gamma_K(x) = \{y \in K : d(x,y) = d_K(x)\}$
- $\theta_K(x)$: center and radius of the smallest ball enclosing $\Gamma_K(x)$

$$\nabla d_K(x) = \frac{x - \theta_K(x)}{d_K(x)}$$

Although not continuous, it can be integrated in a continuous flow.



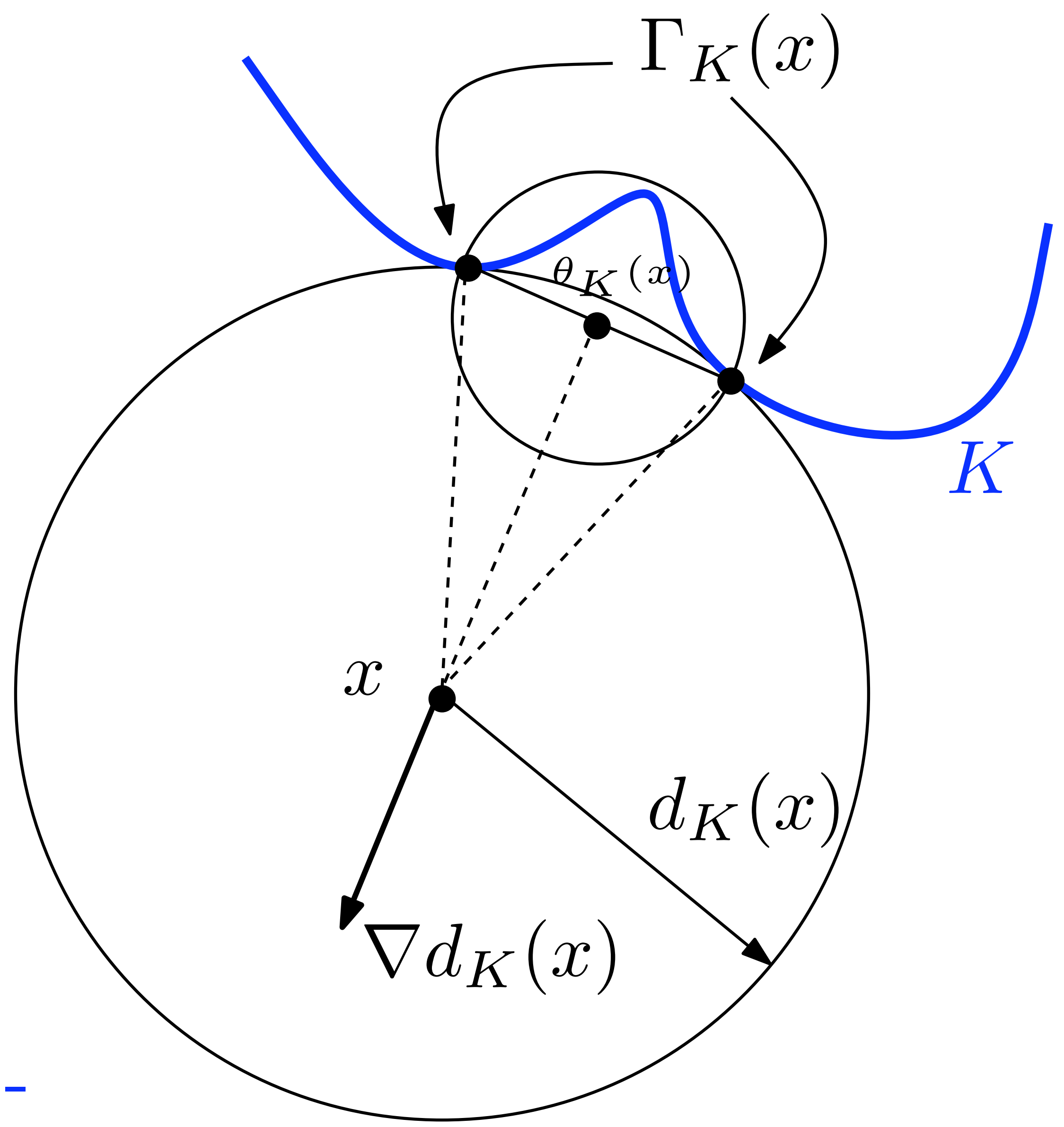
Definition: x is a *critical point* of d_K iff $\nabla d_K(x) = 0$

The gradient of the distance function

- $\Gamma_K(x) = \{y \in K : d(x,y) = d_K(x)\}$
- $\theta_K(x)$: center and radius of the smallest ball enclosing $\Gamma_K(x)$

$$\nabla d_K(x) = \frac{x - \theta_K(x)}{d_K(x)}$$

Can be generalized to distances to compact subsets of complete Riemannian manifolds



Definition: x is a *critical point* of d_K iff $\nabla d_K(x) = 0$

Critical points and offsets topology

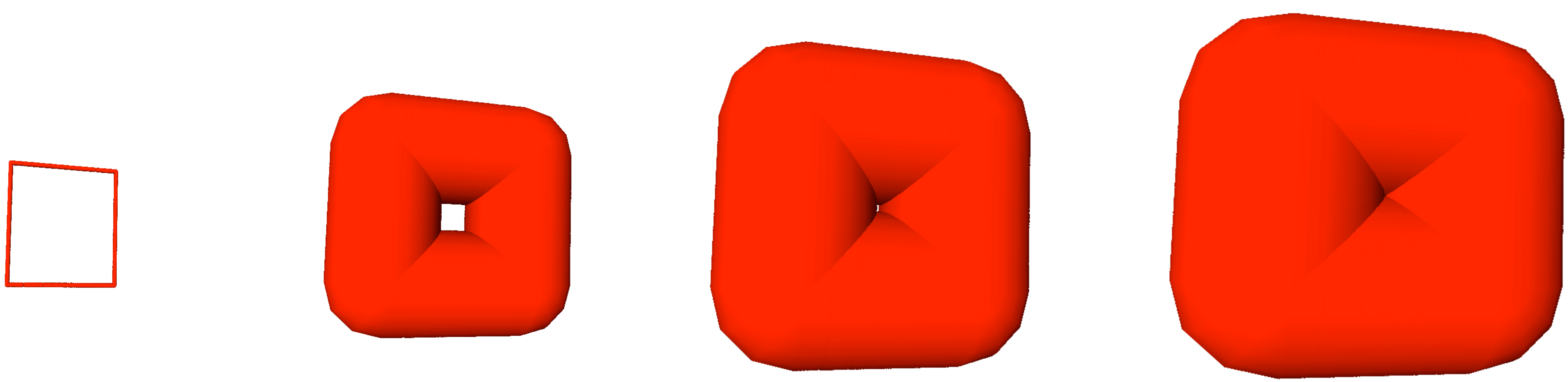
For $\alpha \geq 0$, the *α-offset of K* is $K^\alpha = \{x \in \mathbb{R}^d : d_K(x) \leq \alpha\}$

Theorem: [Grove, Cheeger,...] Let $K \subset \mathbb{R}^d$ be a compact set.

- Let r be a regular value of d_K. Then $d_K^{-1}(r)$ is a topological submanifold of $\mathbb{R}^d$ of codimension 1.
- Let $0 < r_1 < r_2$ be such that $[r_1, r_2]$ does not contain any critical value of d_K. Then all the level sets $d_K^{-1}(r)$, $r \in [r_1, r_2]$ are isotopic and

$$K^{r_2} \setminus K^{r_1} = \{x \in \mathbb{R}^d : r_1 < d_K(x) \leq r_2\}$$

is homeomorphic to $d_K^{-1}(r_1) \times (r_1, r_2]$.



Critical points and offsets topology

For $\alpha \geq 0$, the *α-offset of K* is $K^\alpha = \{x \in \mathbb{R}^d : d_K(x) \leq \alpha\}$

Theorem: [Grove, Cheeger,...] Let $K \subset \mathbb{R}^d$ be a compact set.

- Let r be a regular value of d_K. Then $d_K^{-1}(r)$ is a topological submanifold of $\mathbb{R}^d$ of codimension 1.
- Let $0 < r_1 < r_2$ be such that $[r_1, r_2]$ does not contain any critical value of d_K. Then all the level sets $d_K^{-1}(r)$, $r \in [r_1, r_2]$ are isotopic and

$$K^{r_2} \setminus K^{r_1} = \{x \in \mathbb{R}^d : r_1 < d_K(x) \leq r_2\}$$

  is homeomorphic to $d_K^{-1}(r_1) \times (r_1, r_2]$.

These results still hold for compact sets in (complete) Riemannian manifolds.

Weak feature size and stability

The *weak feature size* of a compact $K \subset \mathbb{R}^d$:

$$\text{wfs}(K) = \inf\{c > 0 : c \text{ is a critical value of } d_K\}$$

Proposition: [C-Lieutier'05] Let $K, K' \subset \mathbb{R}^d$ be such that

$$d_H(K, K') < \varepsilon := \frac{1}{2}\min(\text{wfs}(K), \text{wfs}(K'))$$

Then for all $0 < r \leq 2\varepsilon$, K^r and K'^r are homotopy equivalent.

Weak feature size and stability

The *weak feature size* of a compact $K \subset \mathbb{R}^d$:

$$\text{wfs}(K) = \inf\{c > 0 : c \text{ is a critical value of } d_K\}$$

Proposition: [C-Lieutier'05] Let $K, K' \subset \mathbb{R}^d$ be such that

$$d_H(K, K') < \varepsilon := \frac{1}{2}\min(\text{wfs}(K), \text{wfs}(K'))$$

Then for all $0 < r \leq 2\varepsilon$, K^r and K'^r are homotopy equivalent.

- Two continuous maps $f, f' : X \to Y$ are homotopic if there exist a continuous map $H : X \times [0,1] \to Y$ s.t. $H(.,0) = f$ and $H(.,1) = f'$.
- Two topological spaces X and Y are homotopy equivalent if there exist two continuous maps $f : X \to Y$ and $g : Y \to X$ s. t. $f \circ g$ and $g \circ f$ are homotopic to id_Y and id_X.

Weaker than homeomorphy but "share the same topological invariants" (e.g. Betti numbers)

Weak feature size and stability

The *weak feature size* of a compact $K \subset \mathbb{R}^d$:

$$\mathsf{wfs}(K) = \inf\{c > 0 : c \ \text{ is a critical value of } \ d_K\}$$

Proposition: [C-Lieutier'05] Let $K, K' \subset \mathbb{R}^d$ be such that

$$d_H(K, K') < \varepsilon := \frac{1}{2}\min(\mathsf{wfs}(K), \mathsf{wfs}(K'))$$

Then for all $0 < r \leq 2\varepsilon$, K^r and K'^r are homotopy equivalent.

Proof: Use the gradient vector field (and its flow) to build an explicit homotopy equivalence.

Weak feature size and stability

The *weak feature size* of a compact $K \subset \mathbb{R}^d$:

$$\text{wfs}(K) = \inf\{c > 0 : c \text{ is a critical value of } d_K\}$$

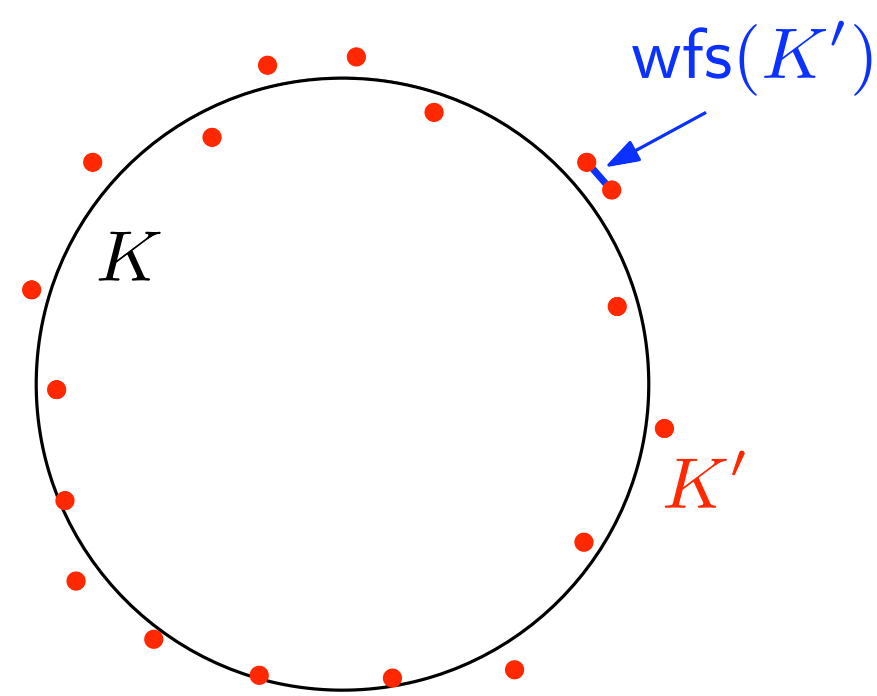
Proposition: [C-Lieutier'05] Let $K, K' \subset \mathbb{R}^d$ be such that

$$d_H(K, K') < \varepsilon := \frac{1}{2} \min(\text{wfs}(K), \text{wfs}(K'))$$

Then for all $0 < r \leq 2\varepsilon$, K^r and K'^r are homotopy equivalent.

Compact set with positive wfs:

- Stability properties
- Large class of compact sets (including sub-analytic sets)
- $K \to \text{wfs}(K)$ is not continuous (unstability of critical points).



Overcoming the discontinuity of wfs

Proposition: [C-Lieutier'05] Let $K, K' \subset \mathbb{R}^d$ be such that

$$d_H(K, K') < \varepsilon := \frac{1}{2} \min(\mathsf{wfs}(K), \mathsf{wfs}(K'))$$

Then for all $0 < r \leq 2\varepsilon$, K^r and K'^r are homotopy equivalent.

$K \to \mathsf{wfs}(K)$ is not continuous (unstability of critical points).

Option 1:
Try to get topological information about K without any assumption on $\mathsf{wfs}(K')$.

Option 2:
Restrict to a smaller class of compact sets with some stability properties of the critical points.

Overcoming the discontinuity of wfs

Proposition: [C-Lieutier'05] Let $K, K' \subset \mathbb{R}^d$ be such that

$$d_H(K, K') < \varepsilon := \frac{1}{2}\min(\mathsf{wfs}(K), \mathsf{wfs}(K'))$$

Then for all $0 < r \leq 2\varepsilon$, K^r and K'^r are homotopy equivalent.

$K \to \mathsf{wfs}(K)$ is not continuous (unstability of critical points).

Option 1:
Try to get topological information about K without any assumption on $\mathsf{wfs}(K')$.

→ Persistence-based inference

Option 2:
Restrict to a smaller class of compact sets with some stability properties of the critical points.

→ Notion of μ-critical points. Strong reconstruction results.

Option 1

Theorem: [C-Lieutier'05-Cohen-Steiner et al '05]
Let $K, K' \subset \mathbb{R}^d$ be compact and let $\varepsilon > 0$ be s.t. $d_H(K, K') < \varepsilon$ and wfs$(K) > 4\varepsilon$. For $\alpha > 0$ s.t. $\alpha + 4\varepsilon <$ wfs(K), let $i : K'^{\alpha+\varepsilon} \hookrightarrow K'^{\alpha+3\varepsilon}$ be the canonical inclusion.For any $0 < r < \mathrm{wfs}(K)$,

$$H_k(K^r) \cong im\left(i_* : H_k(K'^{\alpha+\varepsilon}) \to H_k(K'^{\alpha+3\varepsilon})\right)$$

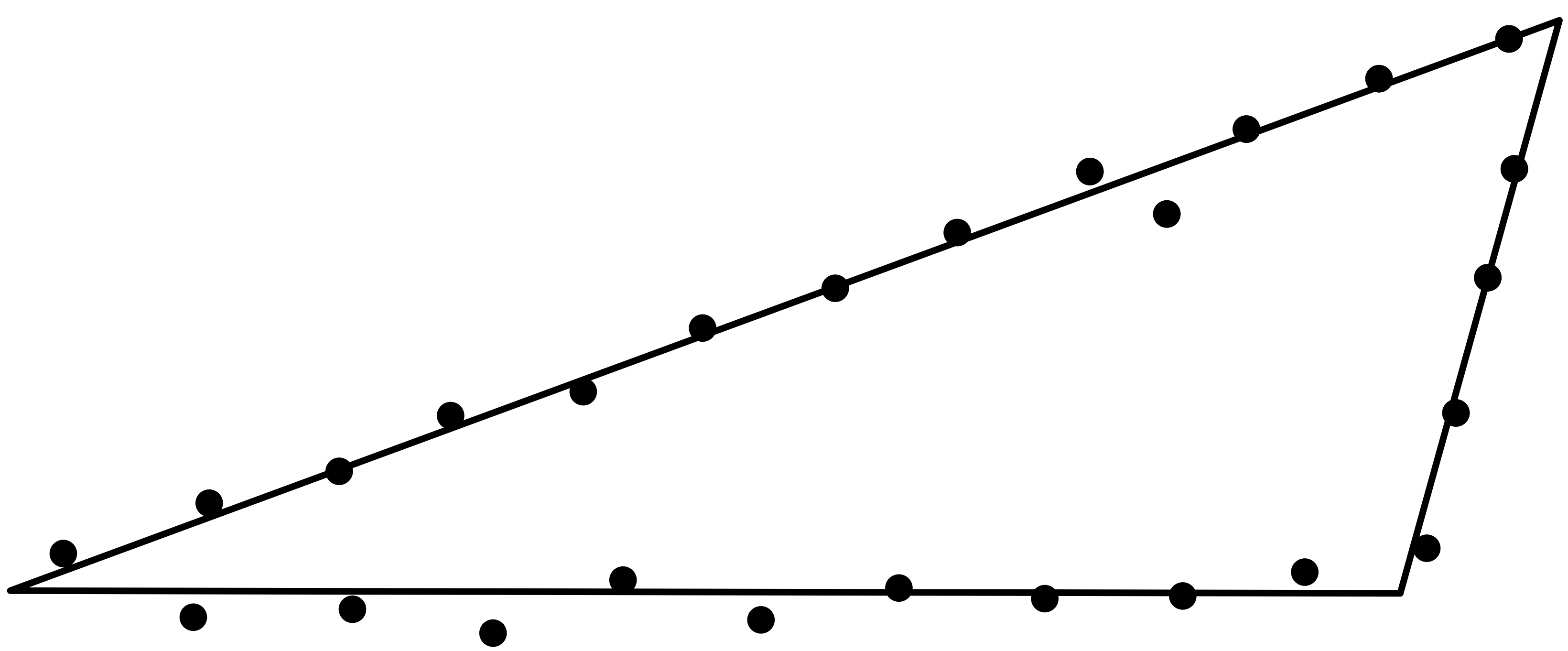
$$\pi_1(K^r, x) \cong im\left(i_* : \pi_1(K'^{\alpha+\varepsilon}, x) \to \pi_1(K'^{\alpha+3\varepsilon}, x)\right)$$

Option 1

Theorem: [C-Lieutier'05-Cohen-Steiner et al '05]
Let $K, K' \subset \mathbb{R}^d$ be compact and let $\varepsilon > 0$ be s.t. $d_H(K, K') < \varepsilon$ and wfs$(K) > 4\varepsilon$. For $\alpha > 0$ s.t. $\alpha + 4\varepsilon <$ wfs(K), let $i : K'^{\alpha+\varepsilon} \hookrightarrow K'^{\alpha+3\varepsilon}$ be the canonical inclusion.For any $0 < r < \mathrm{wfs}(K)$,

$$H_k(K^r) \cong im\left(i_* : H_k(K'^{\alpha+\varepsilon}) \rightarrow H_k(K'^{\alpha+3\varepsilon})\right)$$

$$\pi_1(K^r, x) \cong im\left(i_* : \pi_1(K'^{\alpha+\varepsilon}, x) \rightarrow \pi_1(K'^{\alpha+3\varepsilon}, x)\right)$$

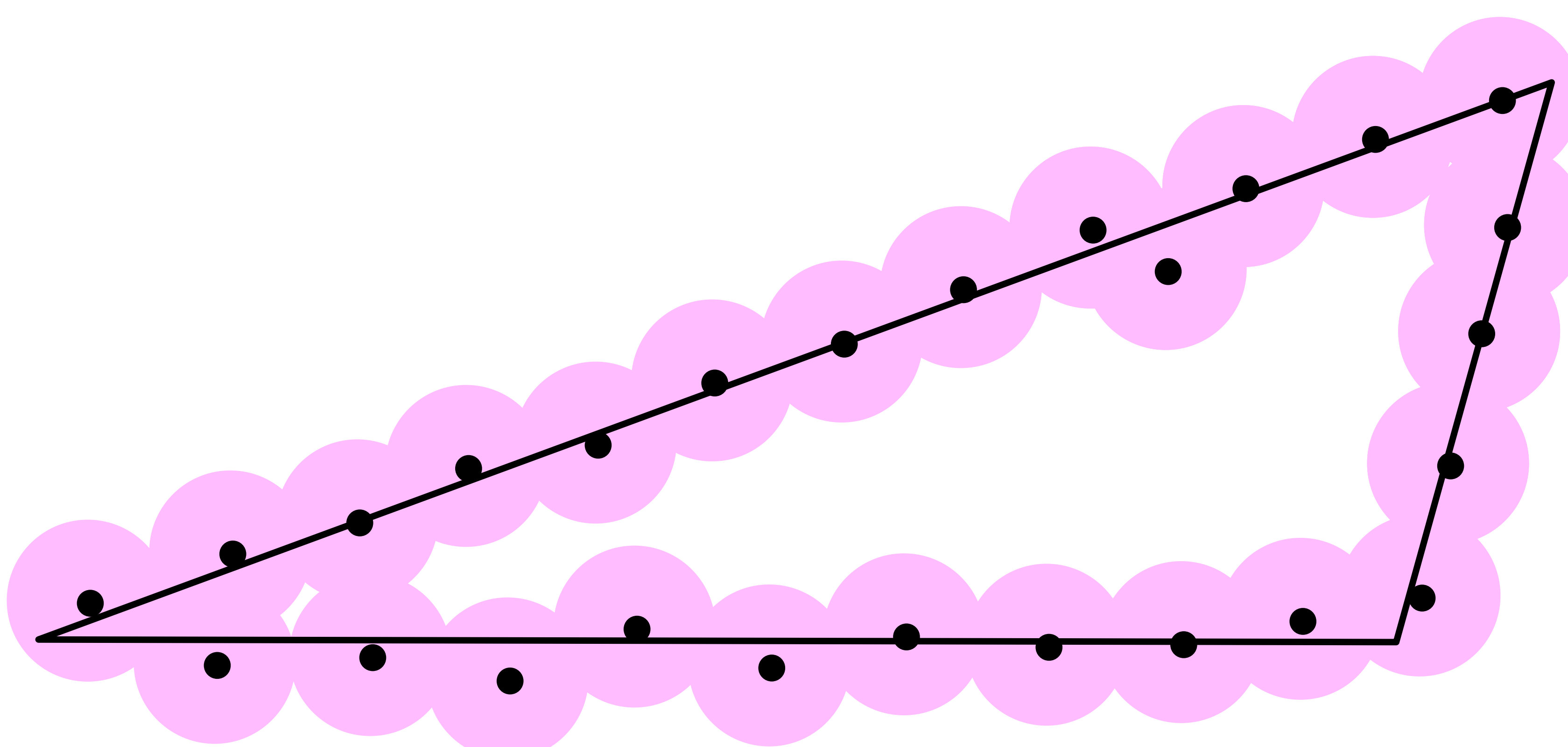


Option 1

Theorem: [C-Lieutier'05-Cohen-Steiner et al '05]
Let $K, K' \subset \mathbb{R}^d$ be compact and let $\varepsilon > 0$ be s.t. $d_H(K, K') < \varepsilon$ and wfs$(K) > 4\varepsilon$. For $\alpha > 0$ s.t. $\alpha + 4\varepsilon <$ wfs(K), let $i : K'^{\alpha+\varepsilon} \hookrightarrow K'^{\alpha+3\varepsilon}$ be the canonical inclusion.For any $0 < r < \mathrm{wfs}(K)$,

$$H_k(K^r) \cong im\left(i_* : H_k(K'^{\alpha+\varepsilon}) \to H_k(K'^{\alpha+3\varepsilon})\right)$$

$$\pi_1(K^r, x) \cong im\left(i_* : \pi_1(K'^{\alpha+\varepsilon}, x) \to \pi_1(K'^{\alpha+3\varepsilon}, x)\right)$$

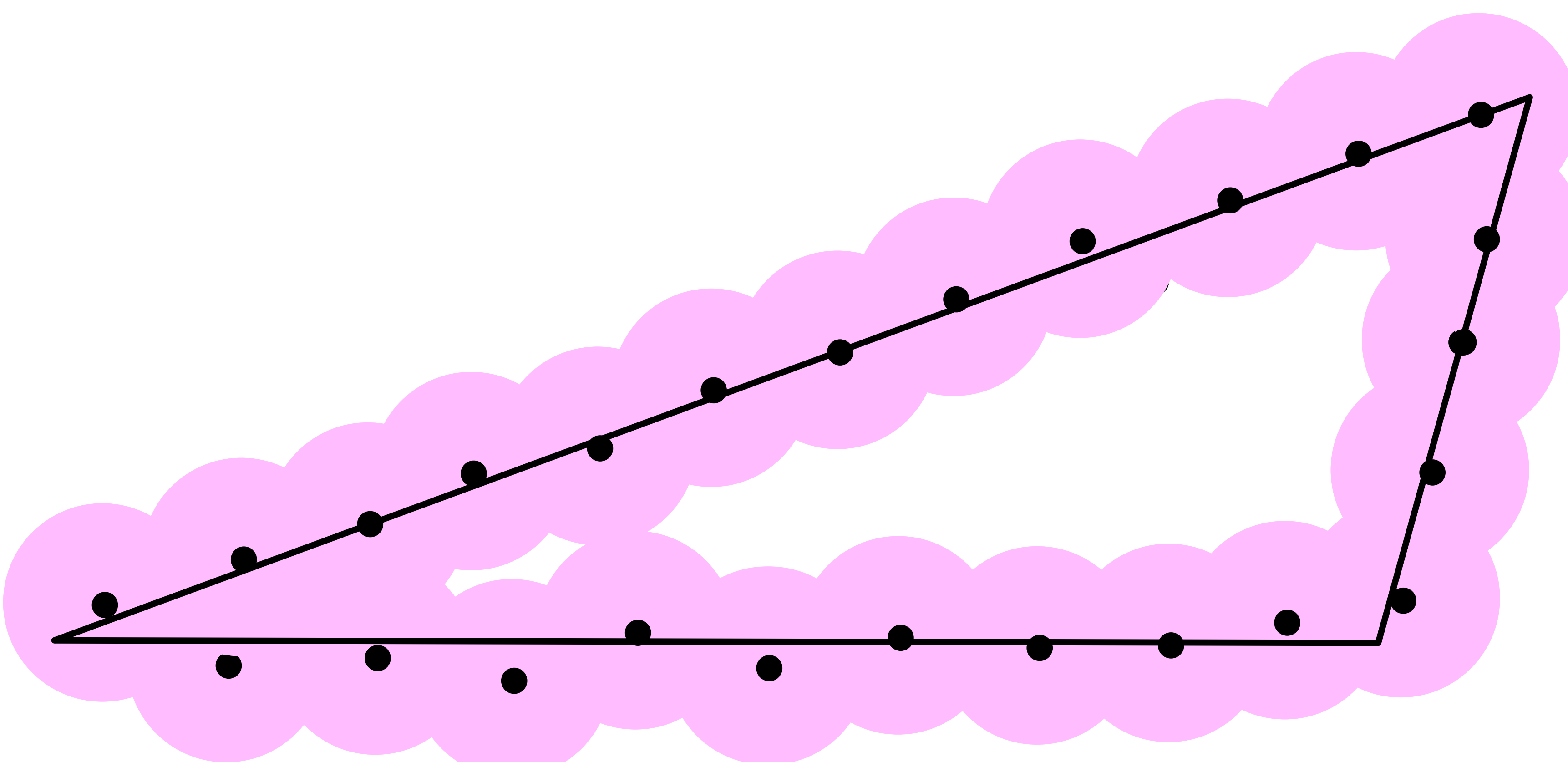


Option 1

Theorem: [C-Lieutier'05-Cohen-Steiner et al '05]
Let $K, K' \subset \mathbb{R}^d$ be compact and let $\varepsilon > 0$ be s.t. $d_H(K, K') < \varepsilon$ and wfs$(K) > 4\varepsilon$. For $\alpha > 0$ s.t. $\alpha + 4\varepsilon <$ wfs(K), let $i : K'^{\alpha+\varepsilon} \hookrightarrow K'^{\alpha+3\varepsilon}$ be the canonical inclusion.For any $0 < r < \mathrm{wfs}(K)$,

$$H_k(K^r) \cong im\left(i_* : H_k(K'^{\alpha+\varepsilon}) \to H_k(K'^{\alpha+3\varepsilon})\right)$$

$$\pi_1(K^r, x) \cong im\left(i_* : \pi_1(K'^{\alpha+\varepsilon}, x) \to \pi_1(K'^{\alpha+3\varepsilon}, x)\right)$$

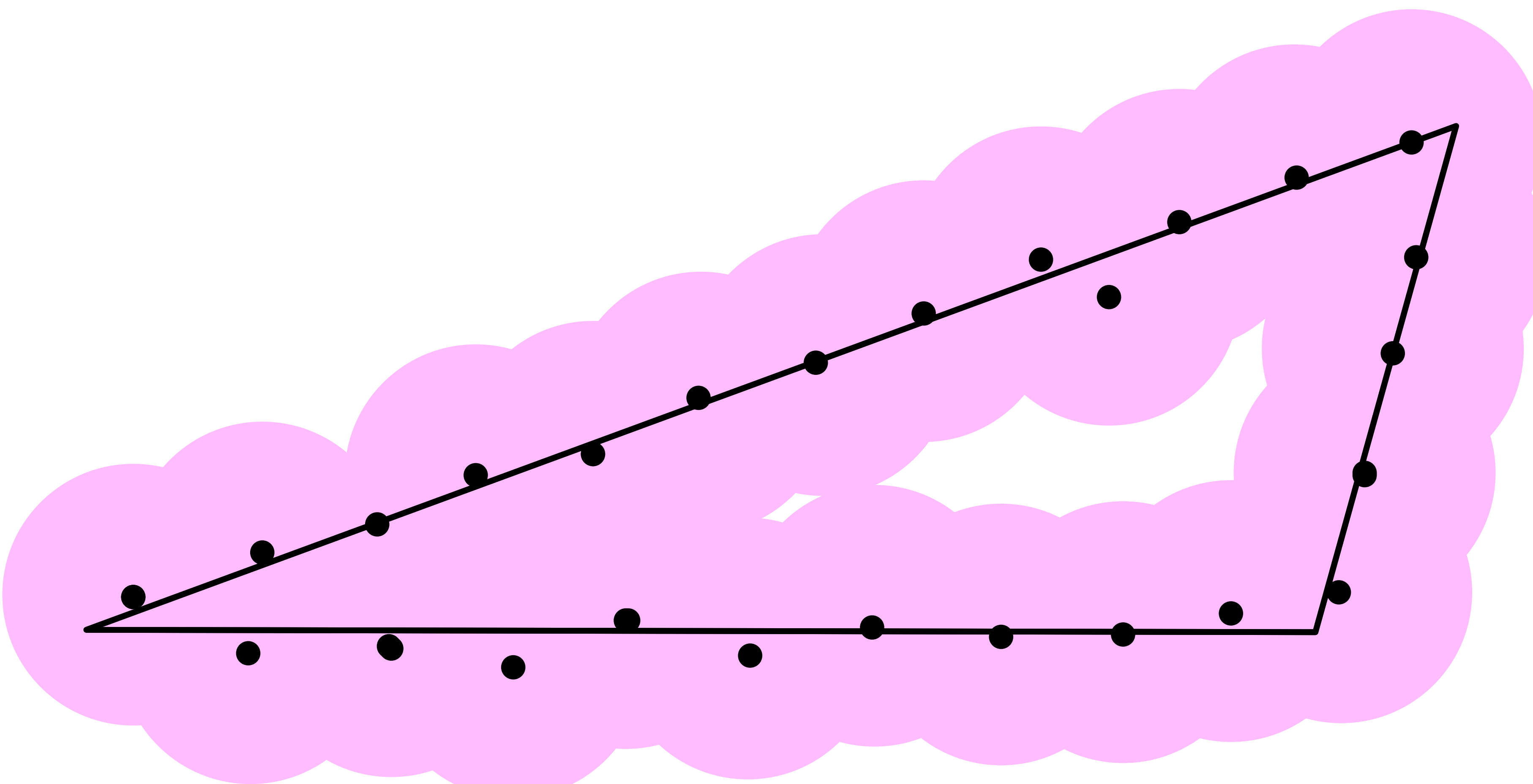


Option 1

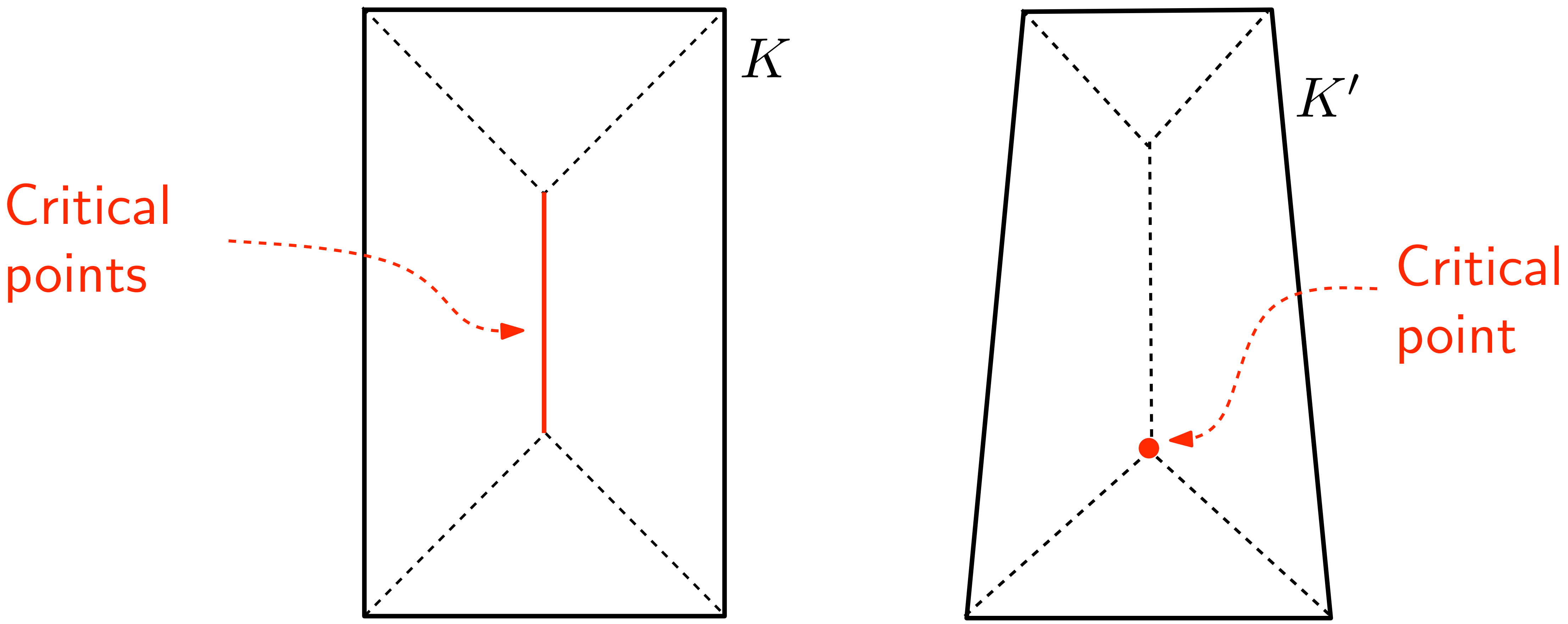
Theorem: [C-Lieutier'05-Cohen-Steiner et al '05]
Let $K, K' \subset \mathbb{R}^d$ be compact and let $\varepsilon > 0$ be s.t. $d_H(K, K') < \varepsilon$ and wfs$(K) > 4\varepsilon$. For $\alpha > 0$ s.t. $\alpha + 4\varepsilon <$ wfs(K), let $i : K'^{\alpha+\varepsilon} \hookrightarrow K'^{\alpha+3\varepsilon}$ be the canonical inclusion.For any $0 < r < \mathrm{wfs}(K)$,

$$H_k(K^r) \cong im\left(i_* : H_k(K'^{\alpha+\varepsilon}) \to H_k(K'^{\alpha+3\varepsilon})\right)$$

$$\pi_1(K^r, x) \cong im\left(i_* : \pi_1(K'^{\alpha+\varepsilon}, x) \to \pi_1(K'^{\alpha+3\varepsilon}, x)\right)$$

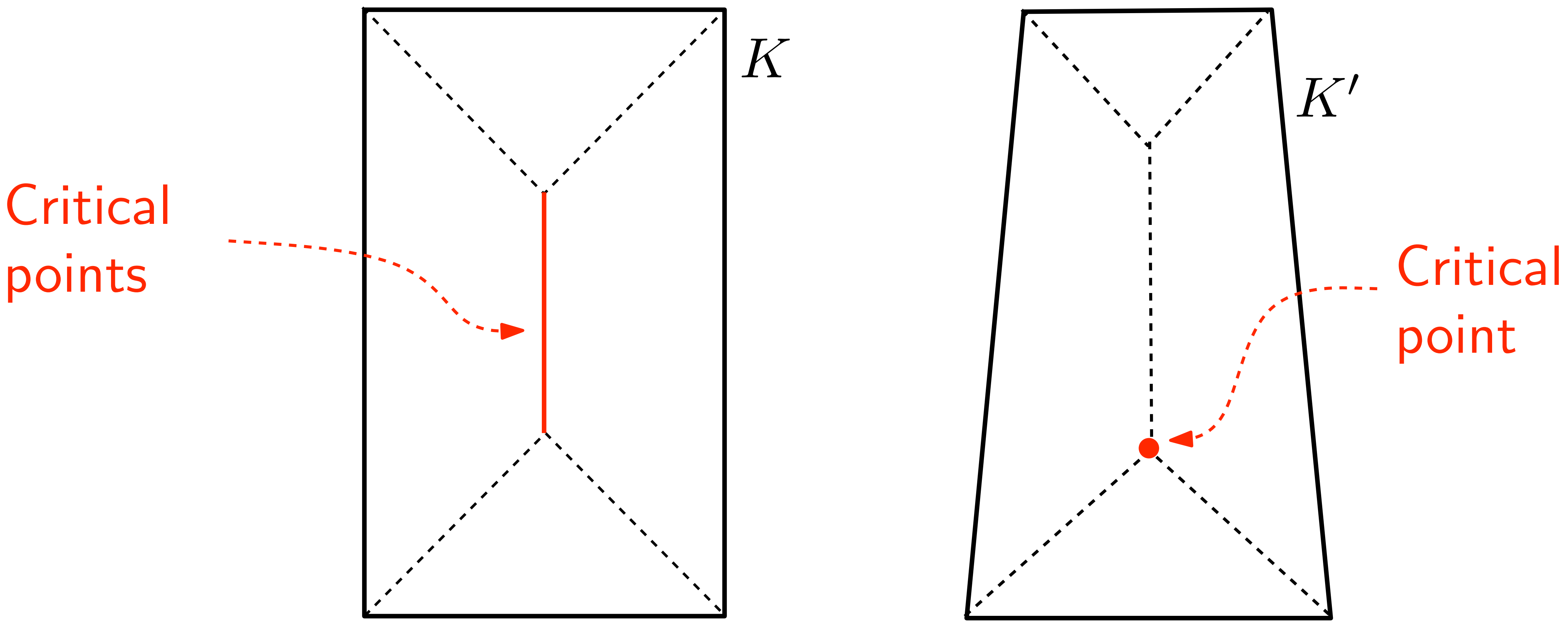


Option 2: μ-critical points and μ-reach



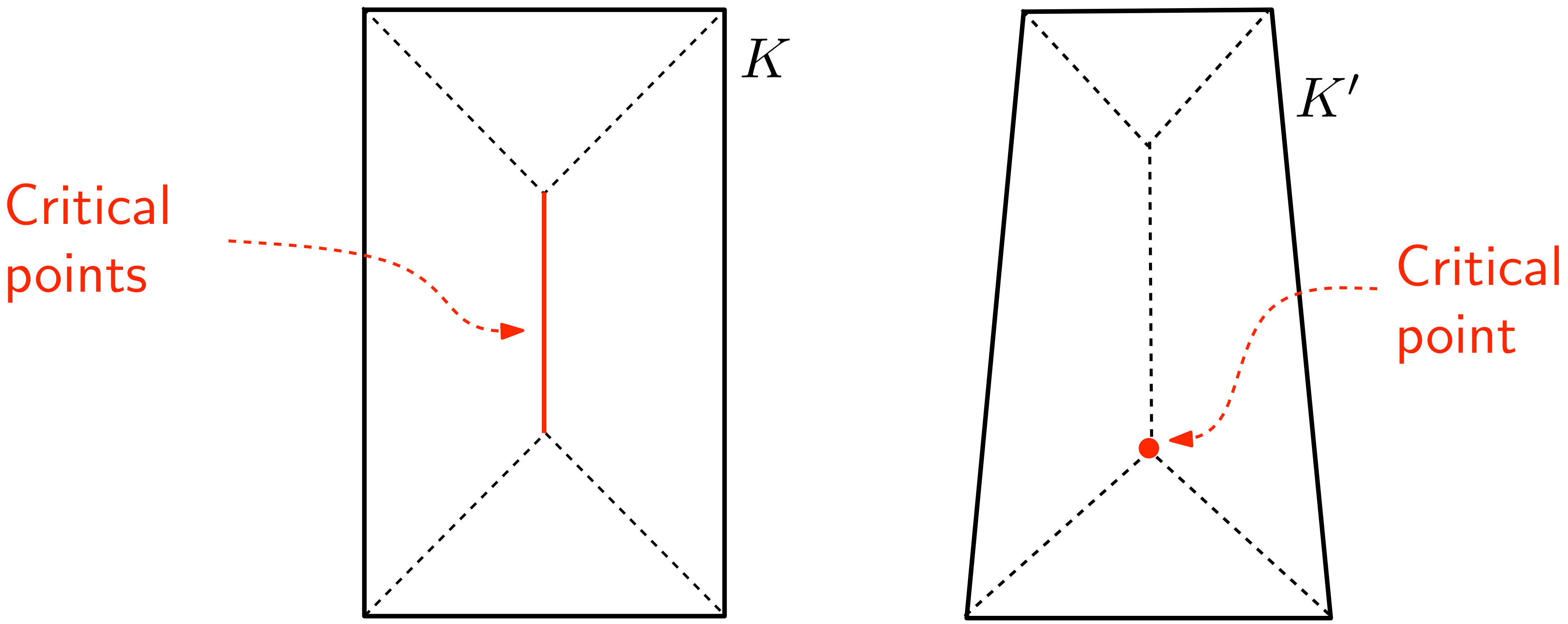
Option 2: μ-critical points and μ-reach

A point $x \in \mathbb{R}^d$ is *μ-critical* for K if $\|\nabla d_K(x)\| \leq \mu$



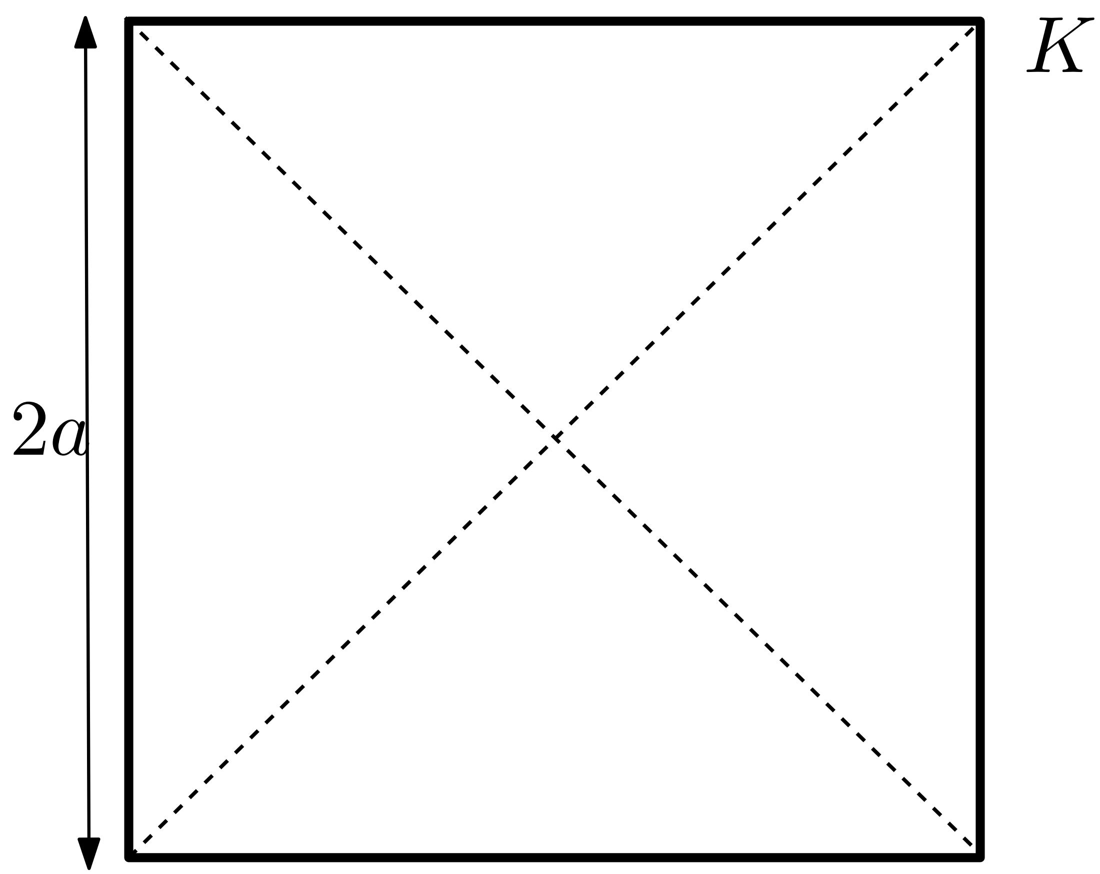
Option 2: μ-critical points and μ-reach

A point $x \in \mathbb{R}^d$ is μ-*critical* for K if $\|\nabla d_K(x)\| \leq \mu$



Theorem: [C-Cohen-Steiner-Lieutier'06] Let $K, K' \subset \mathbb{R}^d$ be two compact sets s. t. $d_H(K, K') \leq \varepsilon$. For any μ-critical point x for K, there exists a $(2\sqrt{\varepsilon / d_K(x)} + \mu)$-critical point for K' at distance at most $2\sqrt{\varepsilon d_K(x)}$ from x.

Option 2: μ-critical points and μ-reach



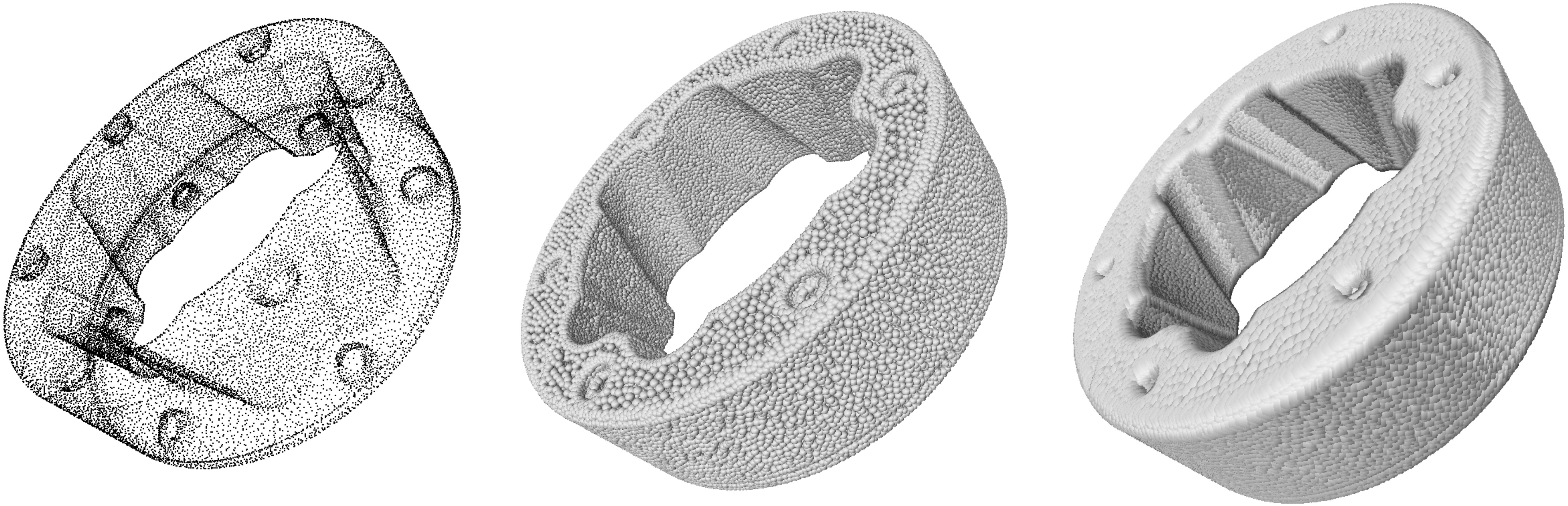
- $r_\mu(K) = 0$ if $\mu \geq \sqrt{2}/2$
- $r_\mu(K) = a$ if $\mu < \sqrt{2}/2$
- $\mathrm{wfs}(K) = a$

μ-*reach* of a compact $K \subset \mathbb{R}^d$:

$$r_\mu(K) = \inf\{d_K(x) : \|\nabla d_K(x)\| < \mu\}$$

- $\forall \mu \in (0, 1)$, $r_\mu(K) \leq$ wfs(K)
- for $\mu = 1$, $r_\mu(K)$ is the reach introduced by Federer in Geometric Measure Theory

Option 2: μ-critical points and μ-reach

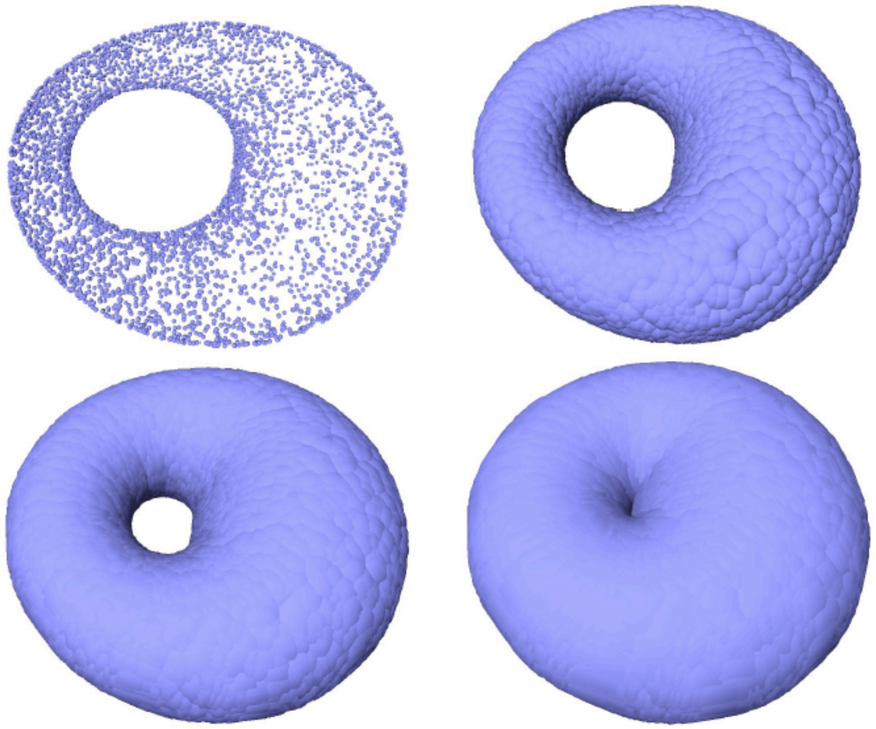


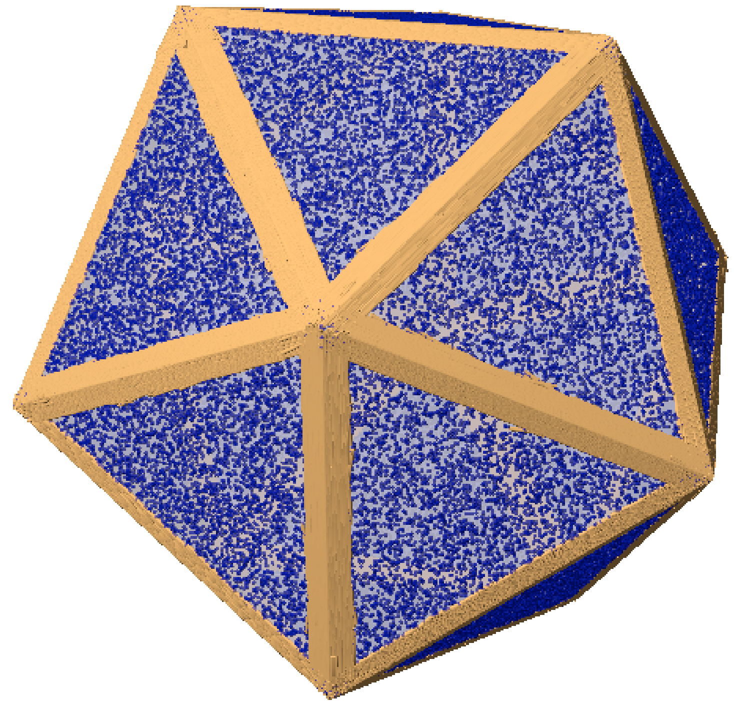
A reconstruction theorem: [C-Cohen-Steiner-Lieutier'06]
Let $K \subset \mathbb{R}^d$ be a compact set s.t. $r_\mu = r_\mu(K) > 0$ for some $\mu > 0$. Let $K \subset \mathbb{R}^d$ be such that $d_H(K, K') < \kappa r_\mu(K)$ with $\kappa < \min(\frac{\sqrt{5}}{2} - 1, \frac{\mu^2}{16+2\mu^2})$
Then for any d, d' s.t.

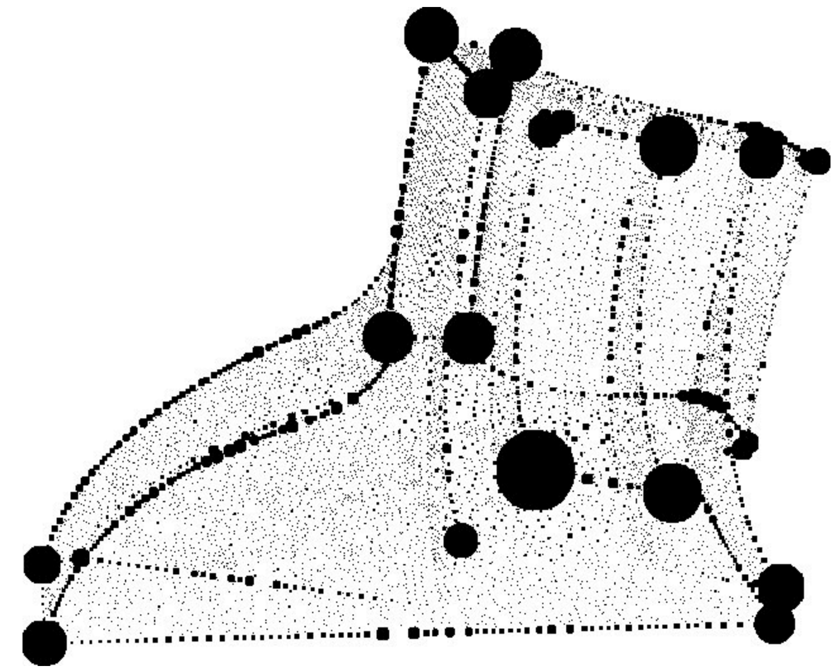
$$0 < d < \text{wfs}(K) \;\; \text{and} \;\; \frac{4\kappa r_\mu}{\mu^2} \leq d' < r_\mu - 3\kappa r_\mu$$

the hypersurfaces $d_{K'}^{-1}(d')$ and $d_K^{-1}(d)$ are isotopic.

Option 2: μ-critical points and μ-reach







Topological/geometric properties of the offsets of K are stable with respect to Hausdorff approximation:

1. Topological stability of the offsets of K (CCSL'06, NSW'06).

2. Approximate normal cones (CCSL'08).

3. Boundary measures (CCSM'07), curvature measures (CCSLT'09), Voronoi covariance measures (GMO'09).

Distance-based inference: the algorithmic side

Topological/geometric inference in practice (from point cloud data sets) ?

Option 2: strong reconstruction results but.....

- Rely on the construction of Voronoï diagram and α-shapes.
- Critical issues in dimension > 3 and non-euclidean spaces.

Option 1:

- Rely on topological persistence theory (at least to infer the homology)
- Efficient algorithms in dimension > 3 and in Riemannian manifolds (or more general metric spaces).

An algorithm for geometric inference

- $X \subset \mathbb{R}^d$ be a compact set such that wfs$(X) > 0$.
- $L \subset \mathbb{R}^d$ be a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon > 0$.

An algorithm for geometric inference

- $X \subset \mathbb{R}^d$ be a compact set such that wfs$(X) > 0$.
- $L \subset \mathbb{R}^d$ be a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon > 0$.

Can be replaced in the following by (complete) Riemannian manifold or a (totally bounded) metric space but require some extra assumptions $\longrightarrow$ see next slides.

An algorithm for geometric inference

- $X \subset \mathbb{R}^d$ be a compact set such that wfs$(X) > 0$.
- $L \subset \mathbb{R}^d$ be a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon > 0$.

Theorem:
Assume that wfs$(X) > 4\varepsilon$. For $\alpha > 0$ s.t. $\alpha + 4\varepsilon <$ wfs(X), let $i : L^{\alpha+\varepsilon} \hookrightarrow L^{\alpha+3\varepsilon}$ be the canonical inclusion.For any $0 < r < \mathrm{wfs}(X)$,

$$H_k(X^r) \cong im\left(i_* : H_k(L^{\alpha+\varepsilon}) \to H_k(L^{\alpha+3\varepsilon})\right)$$

$$\pi_1(X^r, x) \cong im\left(i_* : \pi_1(L^{\alpha+\varepsilon}, x) \to \pi_1(L^{\alpha+3\varepsilon}, x)\right)$$

An algorithm for geometric inference

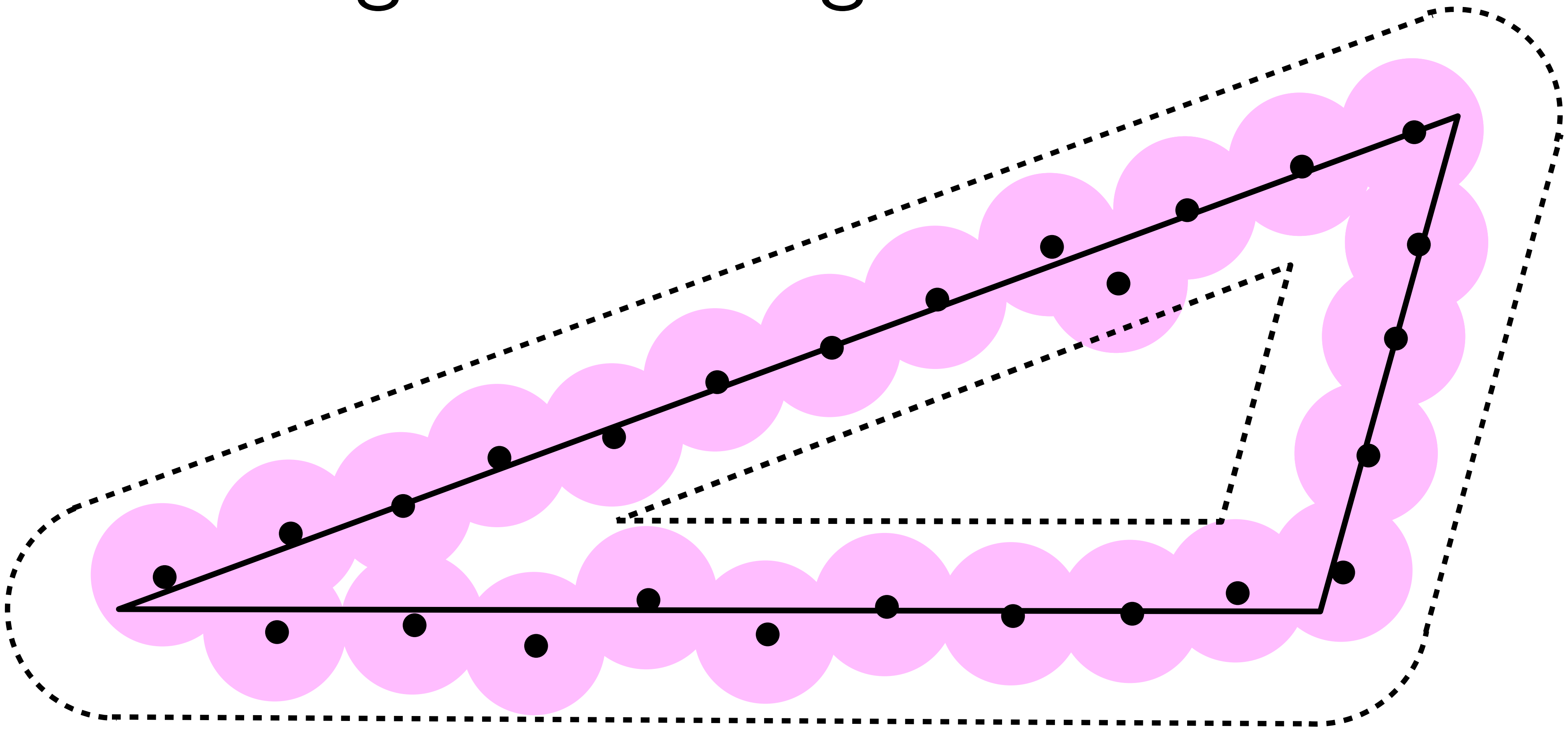
- $X \subset \mathbb{R}^d$ be a compact set such that wfs$(X) > 0$.
- $L \subset \mathbb{R}^d$ be a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon > 0$.

Theorem:
Assume that wfs$(X) > 4\varepsilon$. For $\alpha > 0$ s.t. $\alpha + 4\varepsilon <$ wfs(X), let $i : L^{\alpha+\varepsilon} \hookrightarrow L^{\alpha+3\varepsilon}$ be the canonical inclusion.For any $0 < r < \mathrm{wfs}(X)$,

$$H_k(X^r) \cong im\left(i_* : H_k(L^{\alpha+\varepsilon}) \to H_k(L^{\alpha+3\varepsilon})\right)$$

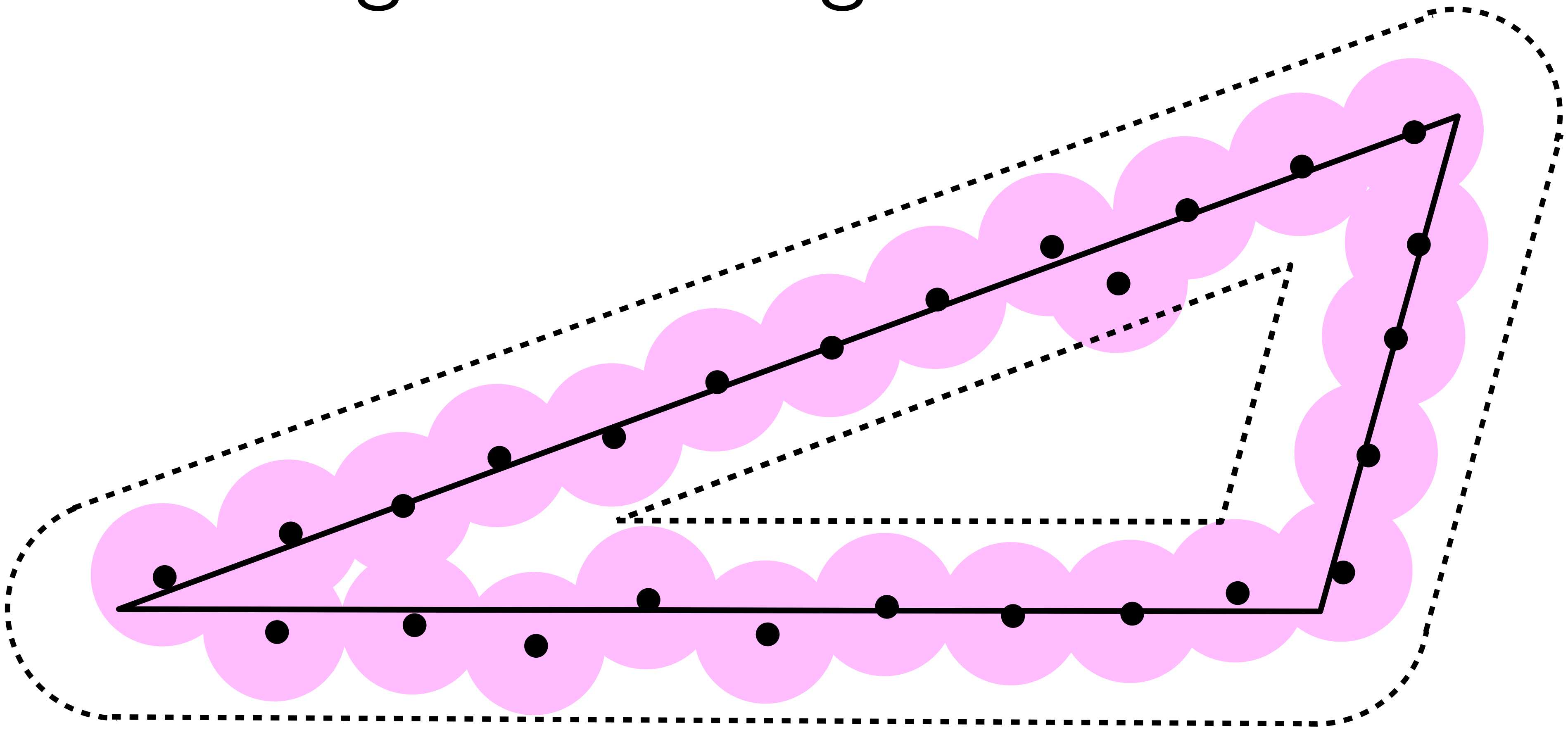
$$\pi_1(X^r, x) \cong im\left(i_* : \pi_1(L^{\alpha+\varepsilon}, x) \to \pi_1(L^{\alpha+3\varepsilon}, x)\right)$$

An algorithm for geometric inference



For any $\alpha > 0$, $\quad X^{\alpha} \subseteq L^{\alpha+\varepsilon} \subseteq X^{\alpha+2\varepsilon} \subseteq L^{\alpha+3\varepsilon} \subseteq X^{\alpha+4\varepsilon} \subseteq \cdots$

An algorithm for geometric inference

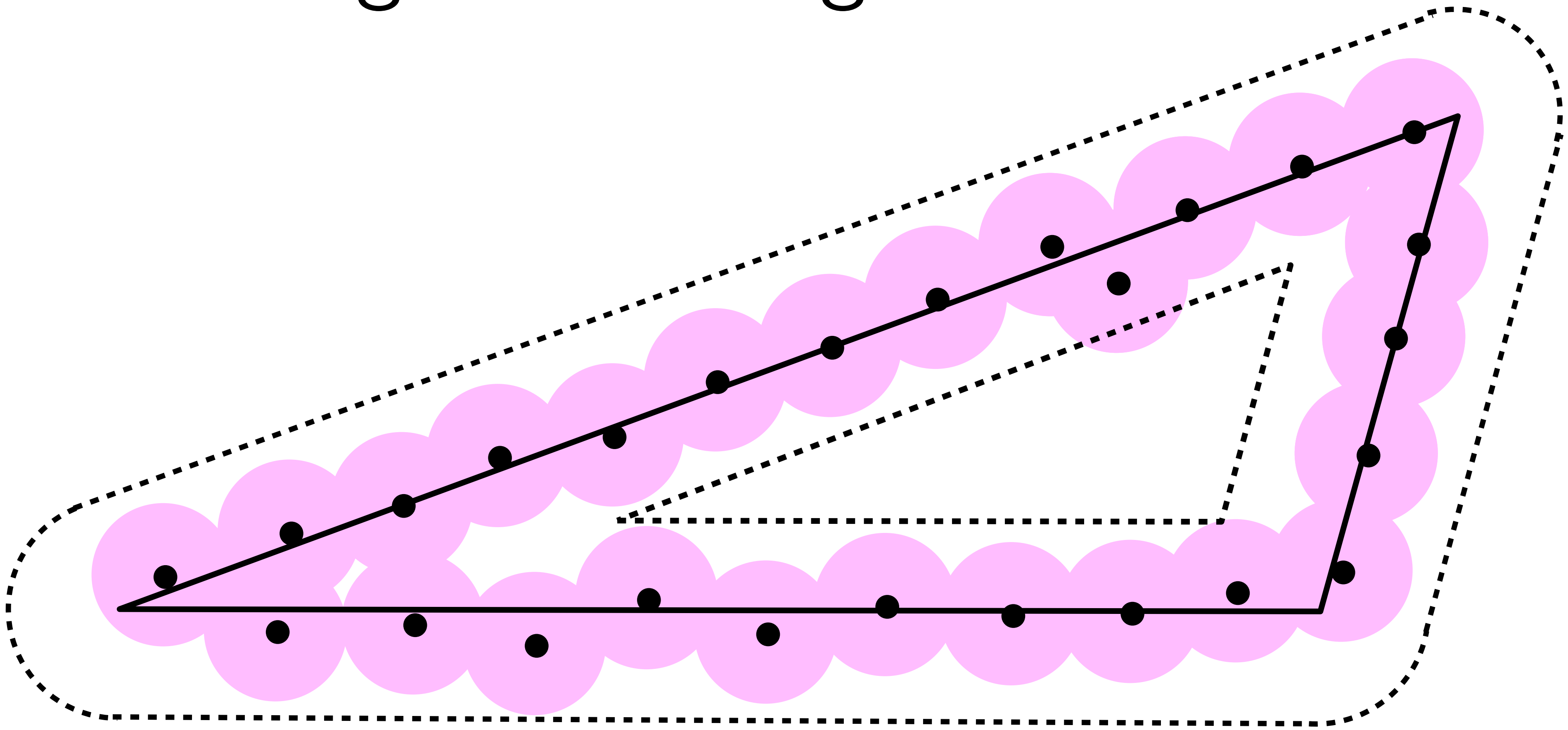


For any $\alpha > 0$, $\quad X^{\alpha} \subseteq L^{\alpha+\varepsilon} \subseteq X^{\alpha+2\varepsilon} \subseteq L^{\alpha+3\varepsilon} \subseteq X^{\alpha+4\varepsilon} \subseteq \cdots$

At homology level:

$$H_k(X^{\alpha}) \to H_k(L^{\alpha+\varepsilon}) \to H_k(X^{\alpha+2\varepsilon}) \to H_k(L^{\alpha+3\varepsilon}) \to H_k(X^{\alpha+4\varepsilon}) \to \cdots$$

An algorithm for geometric inference



For any $\alpha > 0$, $\quad X^{\alpha} \subseteq L^{\alpha+\varepsilon} \subseteq X^{\alpha+2\varepsilon} \subseteq L^{\alpha+3\varepsilon} \subseteq X^{\alpha+4\varepsilon} \subseteq \cdots$

At homology level:

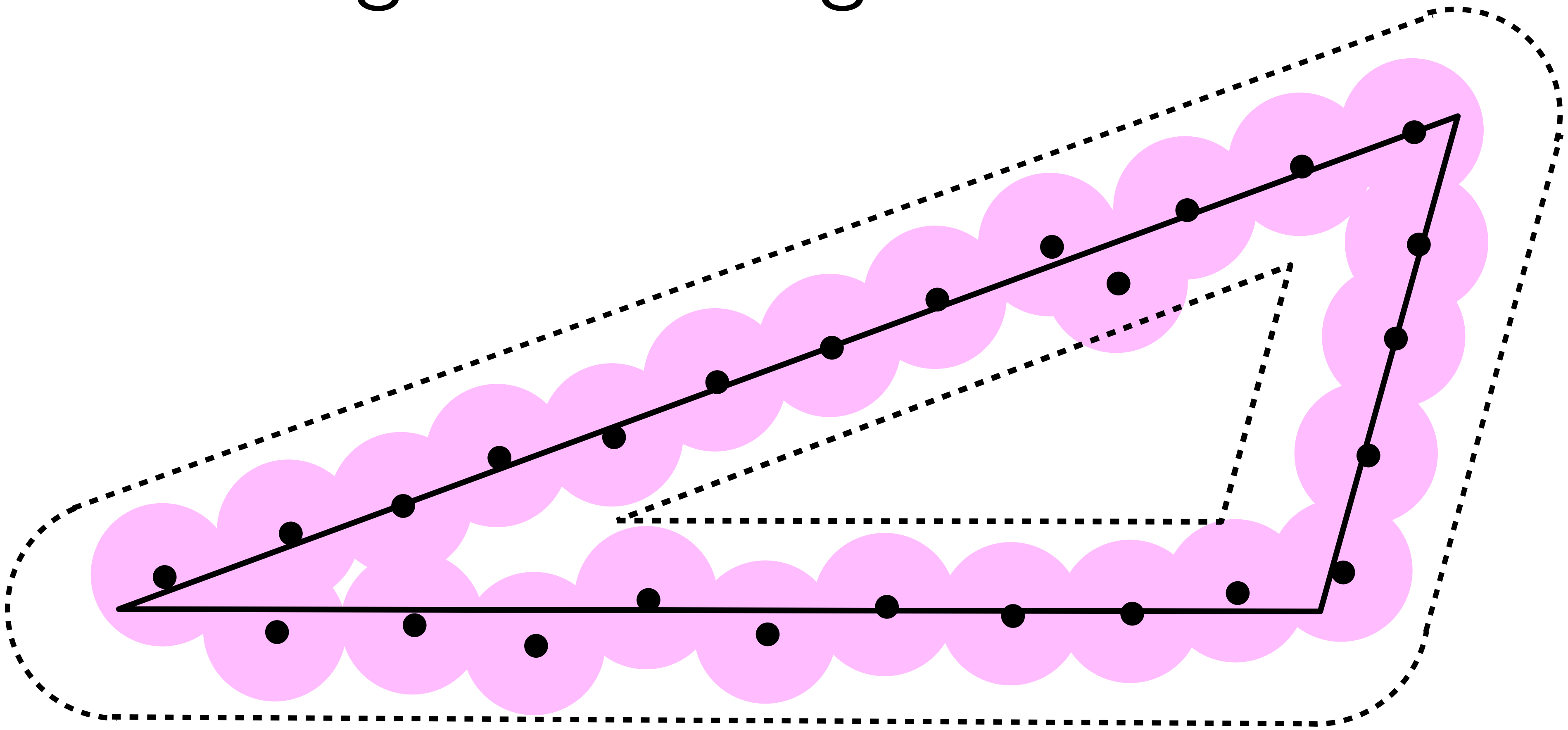
rank $= \dim H_k(X^{\alpha})$

$$H_k(X^{\alpha}) \to H_k(L^{\alpha+\varepsilon}) \to H_k(X^{\alpha+2\varepsilon}) \to H_k(L^{\alpha+3\varepsilon}) \to H_k(X^{\alpha+4\varepsilon}) \to \cdots$$

isomorphism

isomorphism

An algorithm for geometric inference



For any $\alpha > 0$, $\quad X^\alpha \subseteq L^{\alpha+\varepsilon} \subseteq X^{\alpha+2\varepsilon} \subseteq L^{\alpha+3\varepsilon} \subseteq X^{\alpha+4\varepsilon} \subseteq \cdots$

At homology level:

Cannot be directly computed !

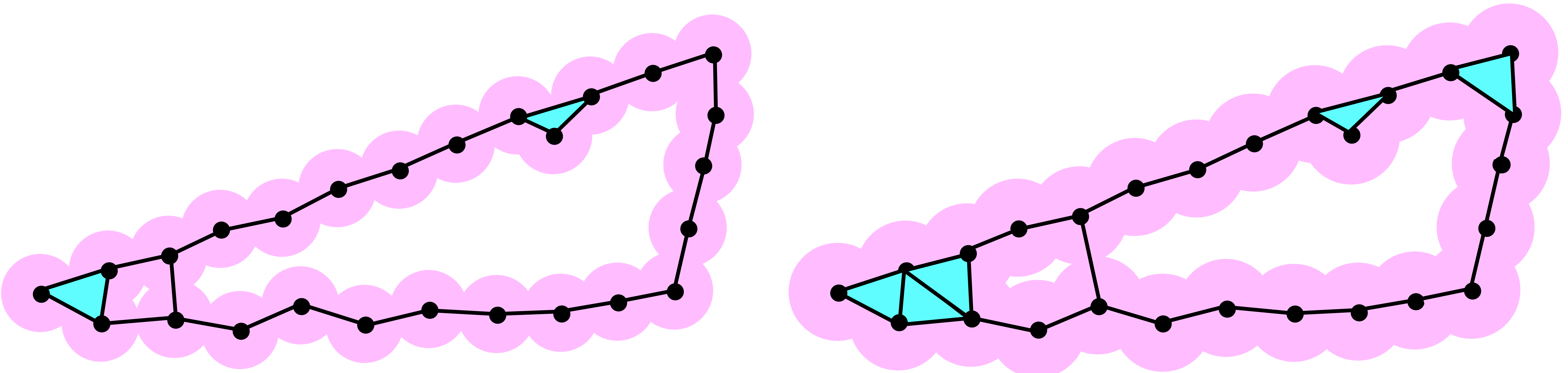
rank $= \dim H_k(X^\alpha)$

$$H_k(X^\alpha) \to H_k(L^{\alpha+\varepsilon}) \to H_k(X^{\alpha+2\varepsilon}) \to H_k(L^{\alpha+3\varepsilon}) \to H_k(X^{\alpha+4\varepsilon}) \to \cdots$$

isomorphism

isomorphism

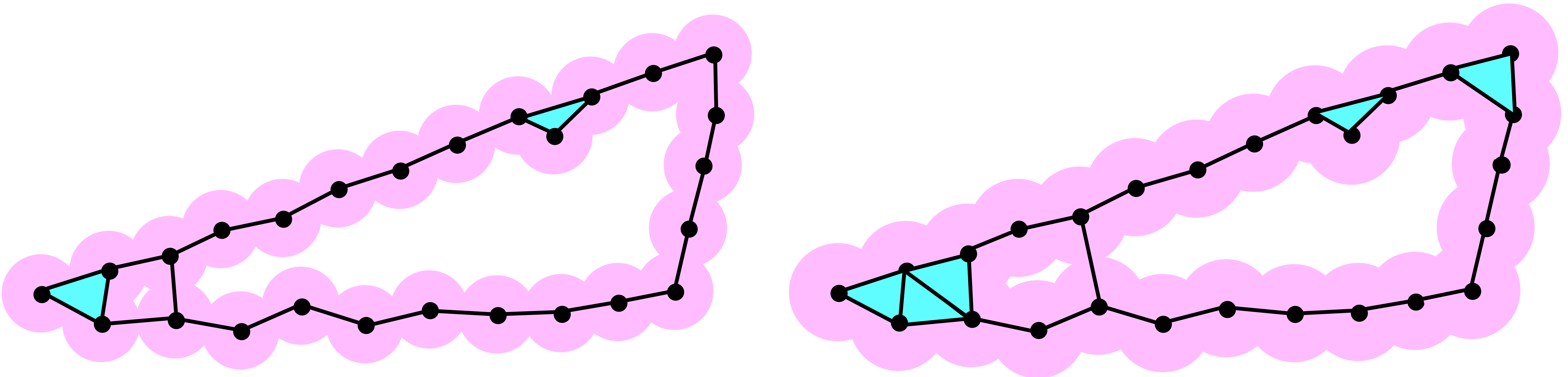
Using the Čech complex



The Čech complex $\mathcal{C}^\alpha(L)$:

$$\text{for } p_0, \cdots p_k \in L, \quad \sigma = [p_0 p_1 \cdots p_k] \in \mathcal{C}^\alpha(L) \;\text{ iff }\; \bigcap_{i=0}^{k} B(p_i, \alpha) \neq \emptyset$$

Using the Čech complex

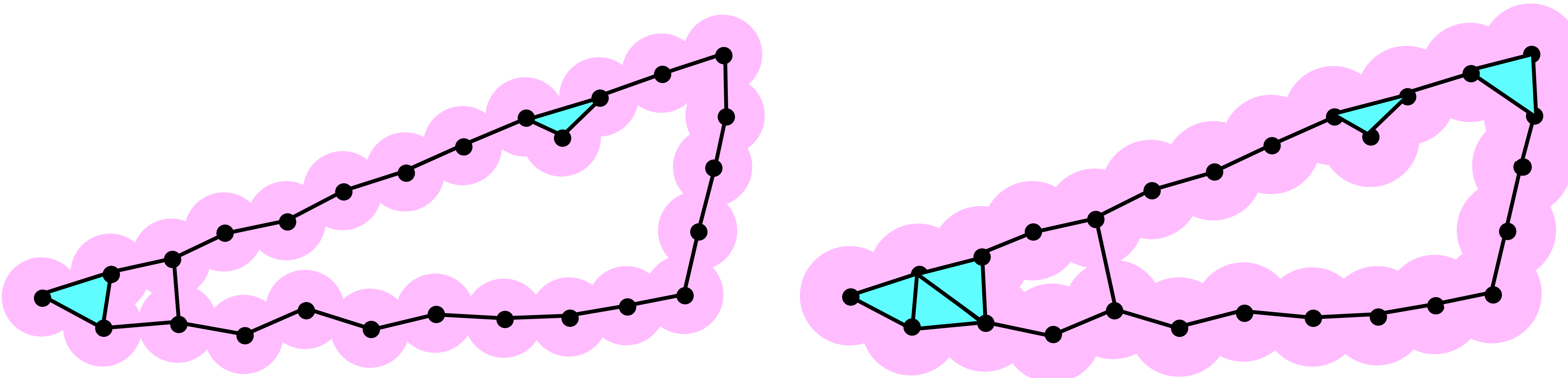


The Čech complex $\mathcal{C}^\alpha(L)$:

$$\text{for } p_0, \cdots p_k \in L, \quad \sigma = [p_0 p_1 \cdots p_k] \in \mathcal{C}^\alpha(L) \text{ iff } \bigcap_{i=0}^{k} B(p_i, \alpha) \neq \emptyset$$

Nerve theorem: For any $\alpha > 0$, L^α and $\mathcal{C}^\alpha(L)$ are homotopy equivalent and the homotopy equivalences can be chosen to commute with inclusions.

Using the Čech complex



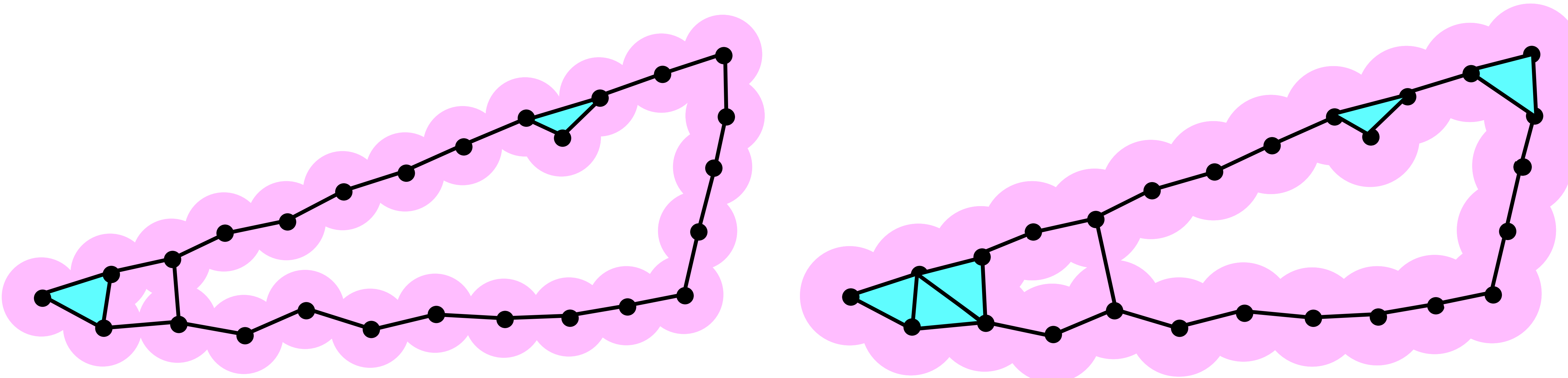
The Čech complex $\mathcal{C}^\alpha(L)$:

$$\text{for } p_0, \cdots p_k \in L, \quad \sigma = [p_0 p_1 \cdots p_k] \in \mathcal{C}^\alpha(L) \;\text{ iff }\; \bigcap_{i=0}^{k} B(p_i, \alpha) \neq \emptyset$$

Nerve theorem: For any $\alpha > 0$, L^α and $\mathcal{C}^\alpha(L)$ are homotopy equivalent and the homotopy equivalences can be chosen to commute with inclusions.

Still true when L is a subset of a Riemannian manifold or a metric space
IF all the intersections $\cap_{i=0}^{k} B(p_i, \alpha)$ are either empty or contractible!

Using the Čech complex



The Čech complex $\mathcal{C}^\alpha(L)$:

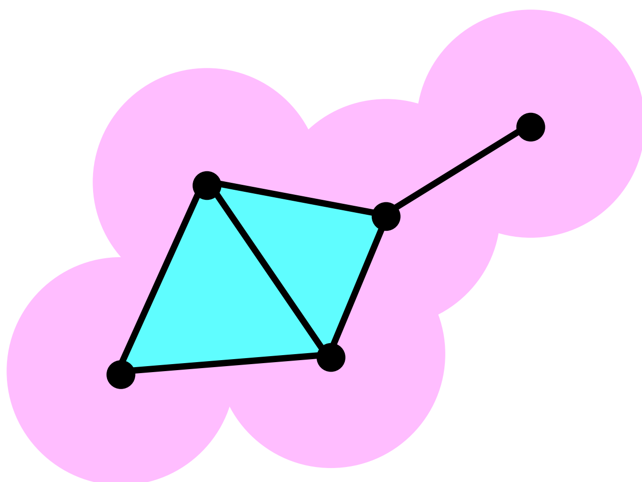
$$\text{for } p_0, \cdots p_k \in L, \quad \sigma = [p_0 p_1 \cdots p_k] \in \mathcal{C}^\alpha(L) \text{ iff } \bigcap_{i=0}^{k} B(p_i, \alpha) \neq \emptyset$$

Nerve theorem: For any $\alpha > 0$, L^α and $\mathcal{C}^\alpha(L)$ are homotopy equivalent and the homotopy equivalences can be chosen to commute with inclusions.

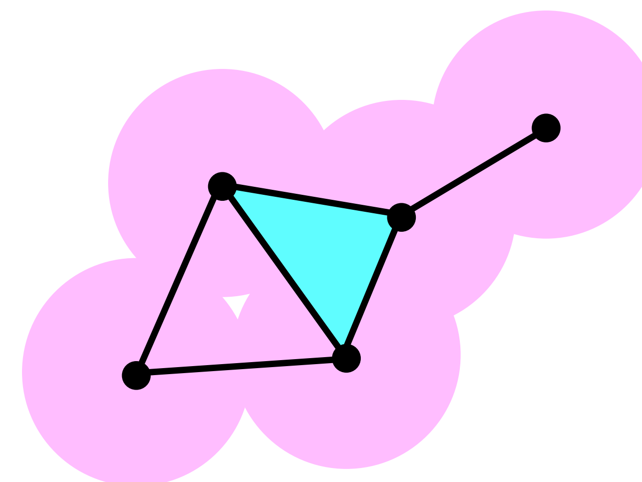
$$\begin{array}{ccccccc} \cdots & \to & H_k(L^{\alpha+\varepsilon}) & \to & H_k(L^{\alpha+3\varepsilon}) & \to & \cdots \\ & & \downarrow & & \downarrow & & \\ \cdots & \to & H_k(\mathcal{C}^{\alpha+\varepsilon}(L)) & \to & H_k(\mathcal{C}^{\alpha+3\varepsilon}(L)) & \to & \cdots \end{array}$$

Allow to work with simplicial complexes but... still too difficult to compute

Using the Rips complex



Rips vs Čech



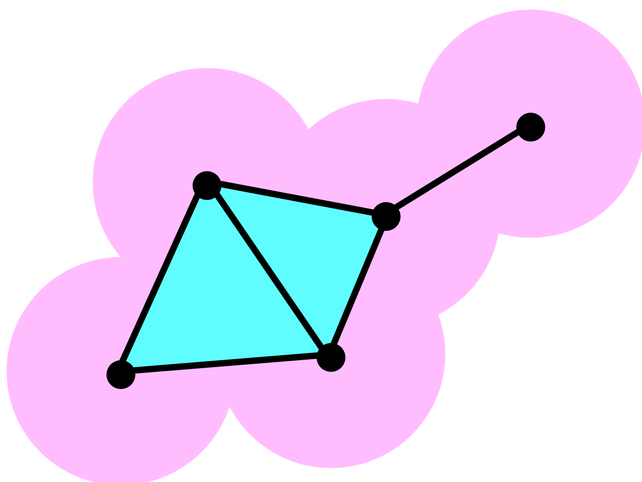
The Rips complex $\mathcal{R}^\alpha(L)$: for $p_0, \cdots p_k \in L$,

$$\sigma = [p_0 p_1 \cdots p_k] \in \mathcal{R}^\alpha(L) \ \text{ iff } \ \forall i, j \in \{0, \cdots k\}, \ d(p_i, p_j) \leq \alpha$$

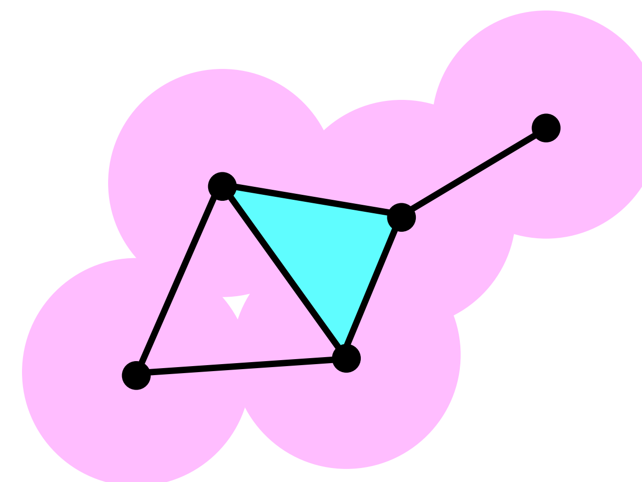
- Easy to compute and fully determined by its 1-skeleton
- Rips-Čech interleaving: for any $\alpha > 0$,

$$\mathcal{C}^{\frac{\alpha}{2}}(L) \subseteq \mathcal{R}^\alpha(L) \subseteq \mathcal{C}^\alpha(L) \subseteq \mathcal{R}^{2\alpha}(L) \subseteq \cdots$$

Using the Rips complex



Rips vs Čech



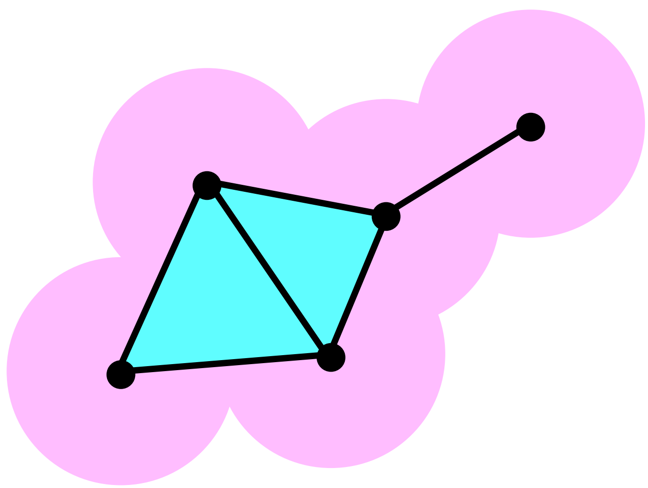
The Rips complex $\mathcal{R}^\alpha(L)$: for $p_0, \cdots p_k \in L$,

$$\sigma = [p_0 p_1 \cdots p_k] \in \mathcal{R}^\alpha(L) \ \text{ iff } \ \forall i,j \in \{0, \cdots k\},\ d(p_i, p_j) \leq \alpha$$

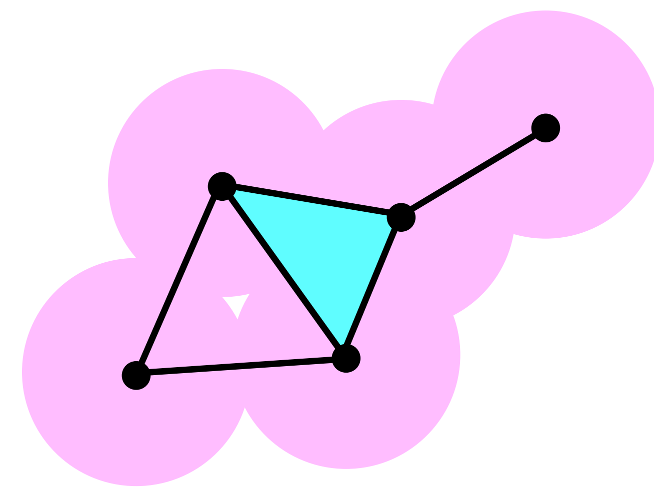
Theorem: [C-Oudot'08]
Let $X \subset \mathbb{R}^d$ be a compact set and $L \subset \mathbb{R}^d$ a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon < \frac{1}{9}$ wfs(X). Then for all $\alpha \in [2\varepsilon, \frac{1}{4}(\mathrm{wfs}(X) - \varepsilon)]$ and all $\lambda \in (0, \mathrm{wfs}(X)))$, one has: $\forall k \in \mathbb{N}$

$$\beta_k(X^\lambda) = \dim(H_k(X^\lambda)) = \mathsf{rk}(\mathcal{R}^\alpha(L) \to \mathcal{R}^{4\alpha}(L))$$

Using the Rips complex



Rips vs Čech



The Rips complex $\mathcal{R}^\alpha(L)$: for $p_0, \cdots p_k \in L$,

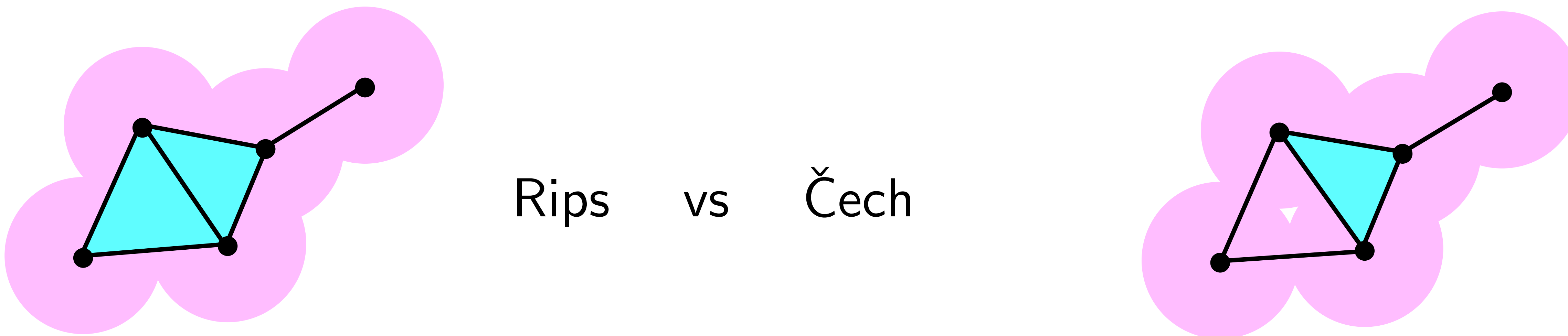
$$\sigma = [p_0 p_1 \cdots p_k] \in \mathcal{R}^\alpha(L) \;\text{ iff }\; \forall i,j \in \{0, \cdots k\},\; d(p_i, p_j) \leq \alpha$$

Theorem: [C-Oudot'08]
Let $X \subset \mathbb{R}^d$ be a compact set and $L \subset \mathbb{R}^d$ a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon < \frac{1}{9}$ wfs(X). Then for all $\alpha \in [2\varepsilon, \frac{1}{4}(\mathrm{wfs}(X) - \varepsilon)]$ and all $\lambda \in (0, \mathrm{wfs}(X))$, one has: $\forall k \in \mathbb{N}$

$$\beta_k(X^\lambda) = \dim(H_k(X^\lambda)) = \boxed{\mathrm{rk}(\mathcal{R}^\alpha(L) \to \mathcal{R}^{4\alpha}(L))}$$

Easy to compute using persistence algo.

Using the Rips complex



The Rips complex $\mathcal{R}^\alpha(L)$: for $p_0, \cdots p_k \in L$,

$$\sigma = [p_0 p_1 \cdots p_k] \in \mathcal{R}^\alpha(L) \text{ iff } \forall i, j \in \{0, \cdots k\},\ d(p_i, p_j) \leq \alpha$$

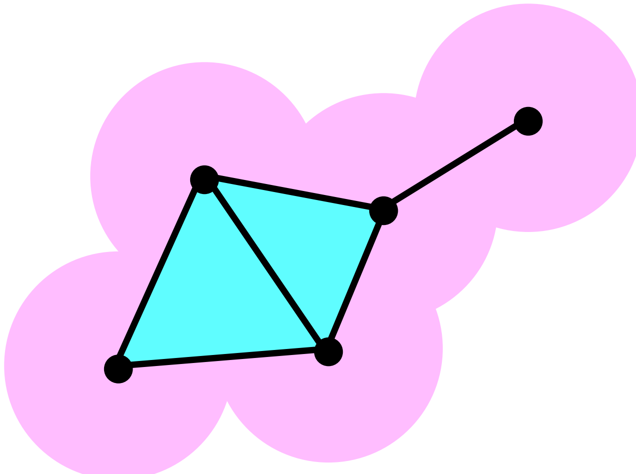
Theorem: [C-Oudot'08]
Let $X \subset \mathbb{R}^d$ be a compact set and $L \subset \mathbb{R}^d$ a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon < \frac{1}{9}$ wfs(X). Then for all $\alpha \in [2\varepsilon, \frac{1}{4}(\text{wfs}(X) - \varepsilon)]$ and all $\lambda \in (0, \text{wfs}(X))$, one has: $\forall k \in \mathbb{N}$

$$\beta_k(X^\lambda) = \dim(H_k(X^\lambda)) = \text{rk}(\mathcal{R}^\alpha(L) \to \mathcal{R}^{4\alpha}(L))$$

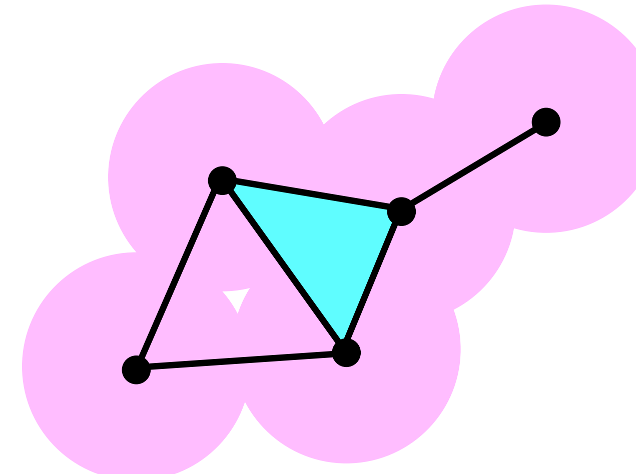
Can be replace by a Riemmanian manifold
BUT take care of convexity radius!
Also some stability results in metric spaces...

Easy to compute using persistence algo.

Using the Rips complex



Rips vs Čech



The Rips complex $\mathcal{R}^\alpha(L)$: for $p_0, \cdots p_k \in L$,

$$\sigma = [p_0 p_1 \cdots p_k] \in \mathcal{R}^\alpha(L) \text{ iff } \forall i, j \in \{0, \cdots k\},\ d(p_i, p_j) \leq \alpha$$

Theorem: [C-Oudot'08]
Let $X \subset \mathbb{R}^d$ be a compact set and $L \subset \mathbb{R}^d$ a finite set such that $d_H(X, L) < \varepsilon$ for some $\varepsilon < \frac{1}{9}$ wfs(X). Then for all $\alpha \in [2\varepsilon, \frac{1}{4}(\mathrm{wfs}(X) - \varepsilon)]$ and all $\lambda \in (0, \mathrm{wfs}(X))$, one has: $\forall k \in \mathbb{N}$

$$\beta_k(X^\lambda) = \dim(H_k(X^\lambda)) = \mathrm{rk}(\mathcal{R}^\alpha(L) \rightarrow \mathcal{R}^{4\alpha}(L))$$

Pb: Choice of α when wfs(X) is unknown?

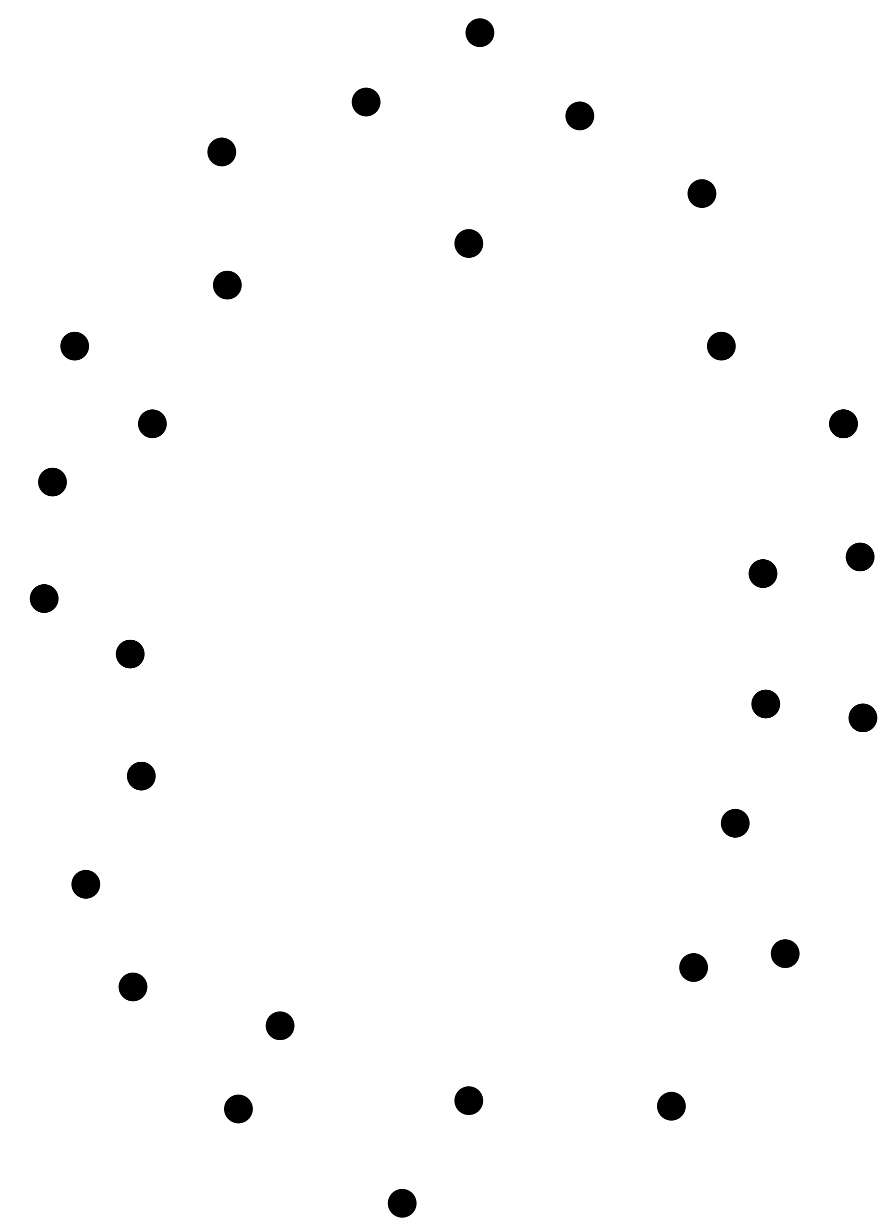
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w'\in W}$.
$\rightarrow$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

Init.: $L = \emptyset$; $\varepsilon = +\infty$

WHILE $L \subset W$
insert $p = argmax_{w\in W} d(w, L)$ in L
update $\varepsilon = \max_{w\in W} d(w, L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers
of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



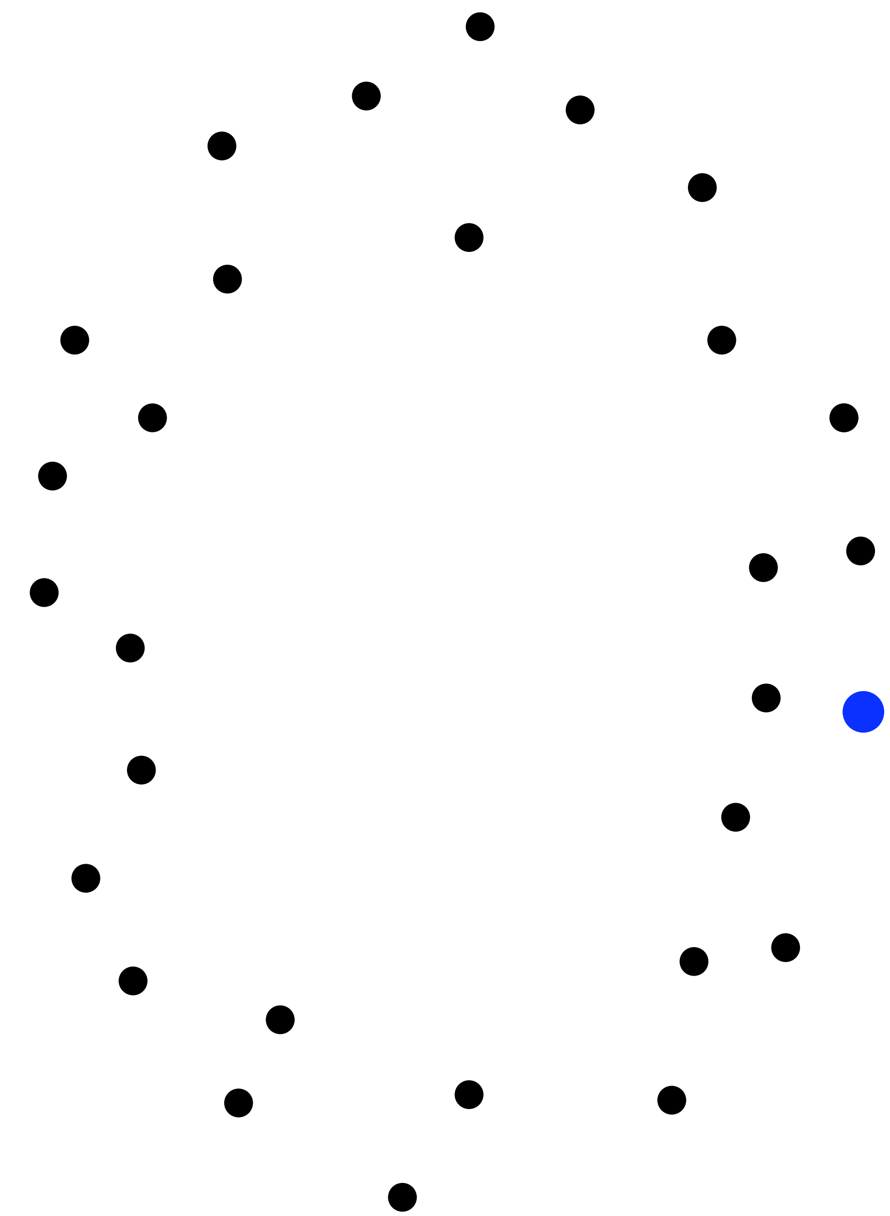
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w'\in W}$.
$\rightarrow$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

Init.: $L = \emptyset$; $\varepsilon = +\infty$

WHILE $L \subset W$
insert $p = argmax_{w\in W} d(w,L)$ in L
update $\varepsilon = \max_{w\in W} d(w,L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



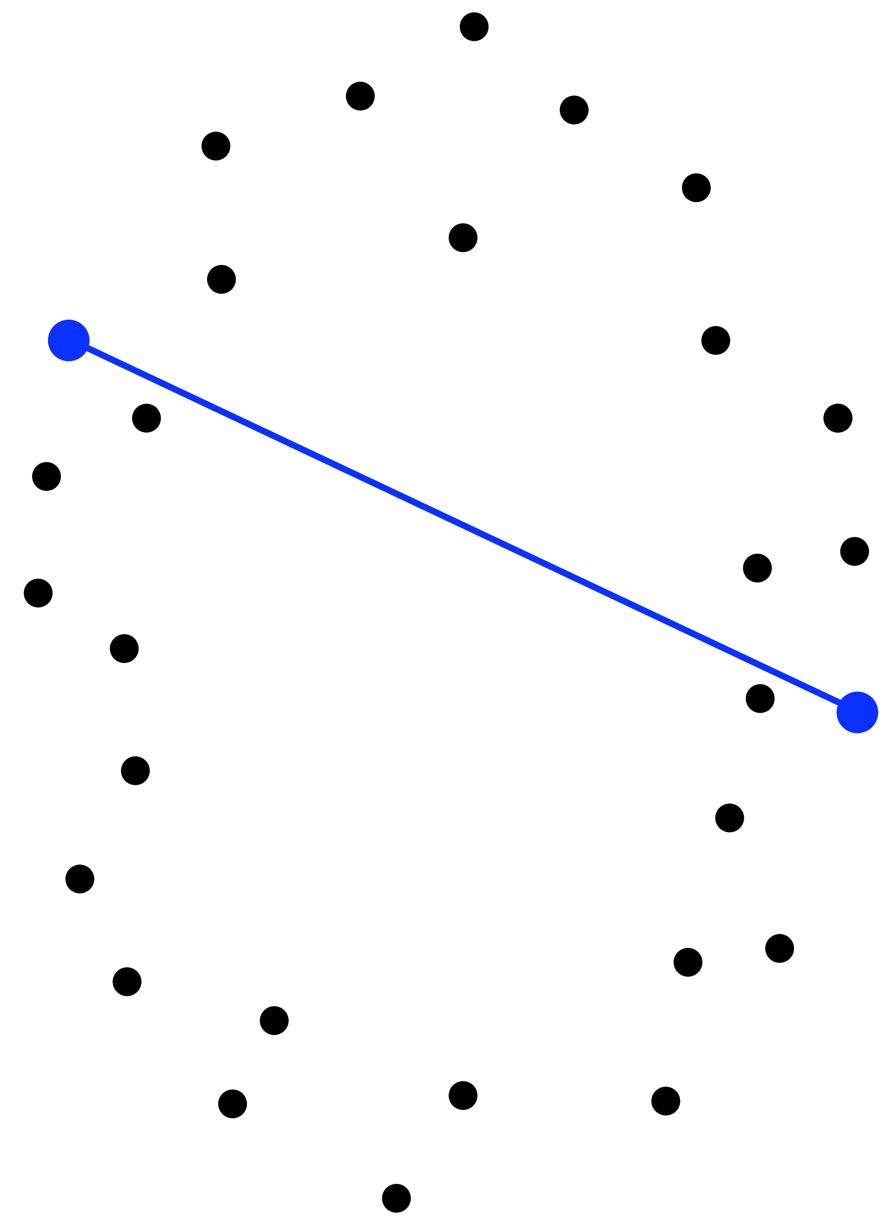
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w'\in W}$.
$\rightarrow$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

Init.: $L = \emptyset$; $\varepsilon = +\infty$

WHILE $L \subset W$
insert $p = argmax_{w\in W} d(w,L)$ in L
update $\varepsilon = \max_{w\in W} d(w,L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



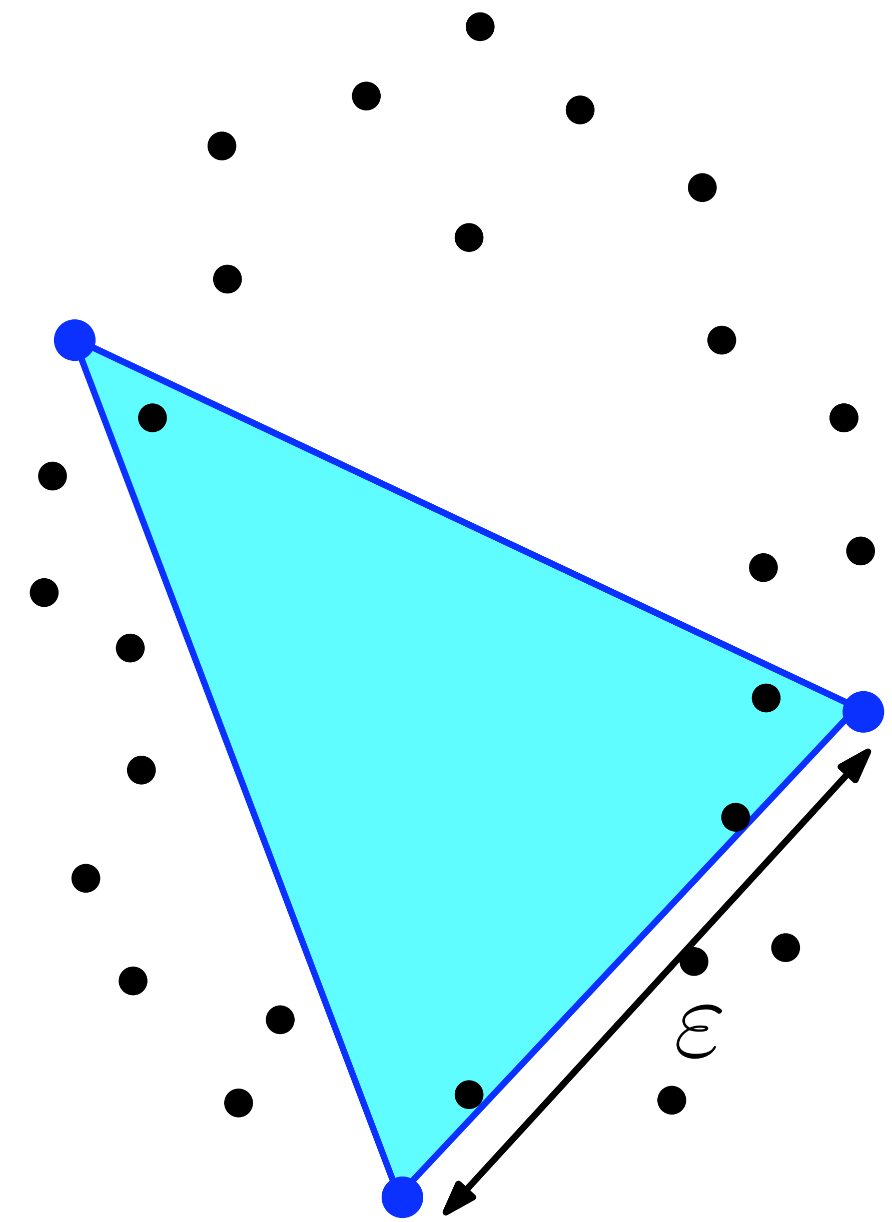
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w'\in W}$.
$\to$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

Init.: $L = \emptyset$; $\varepsilon = +\infty$

WHILE $L \subset W$
insert $p = argmax_{w\in W} d(w,L)$ in L
update $\varepsilon = \max_{w\in W} d(w,L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



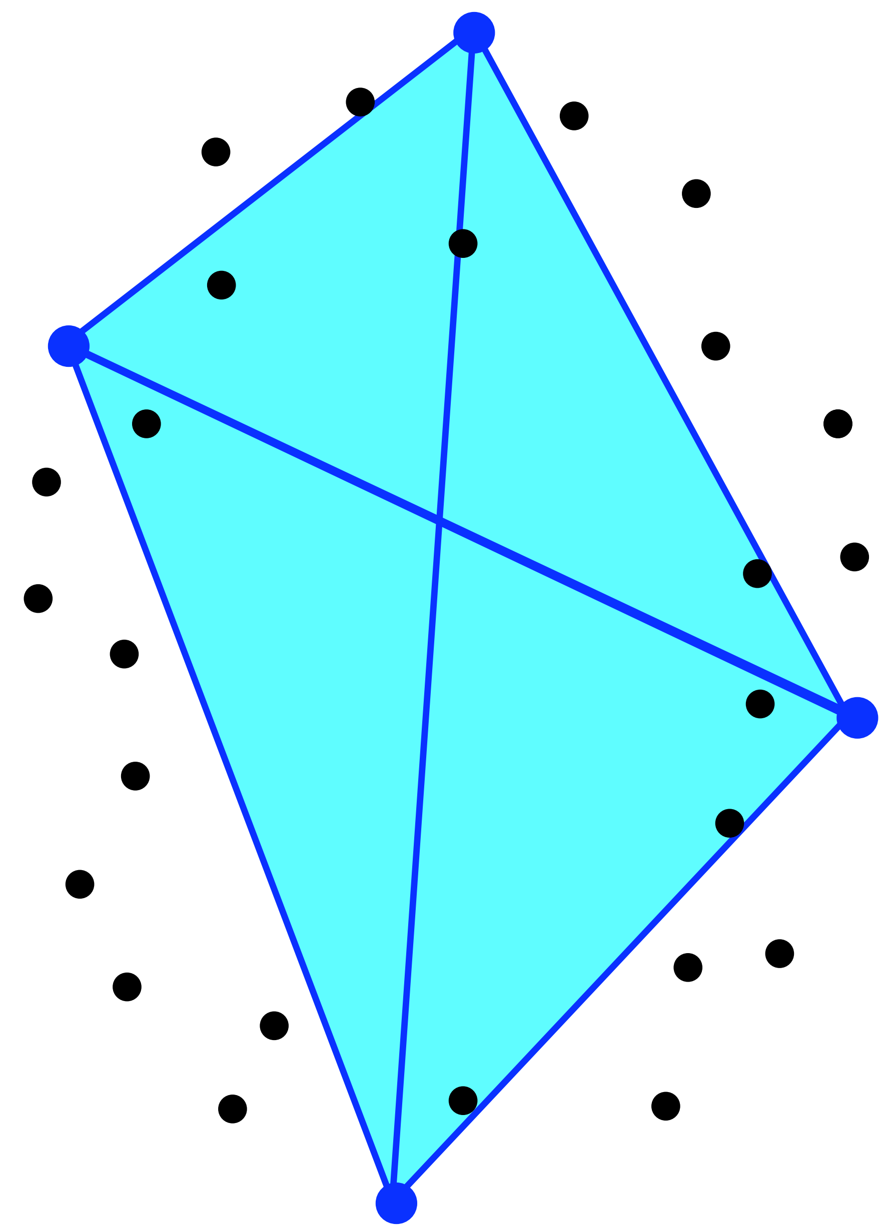
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w'\in W}$.
$\longrightarrow$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

Init.: $L = \emptyset$; $\varepsilon = +\infty$

WHILE $L \subset W$
insert $p = argmax_{w\in W} d(w,L)$ in L
update $\varepsilon = \max_{w\in W} d(w,L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



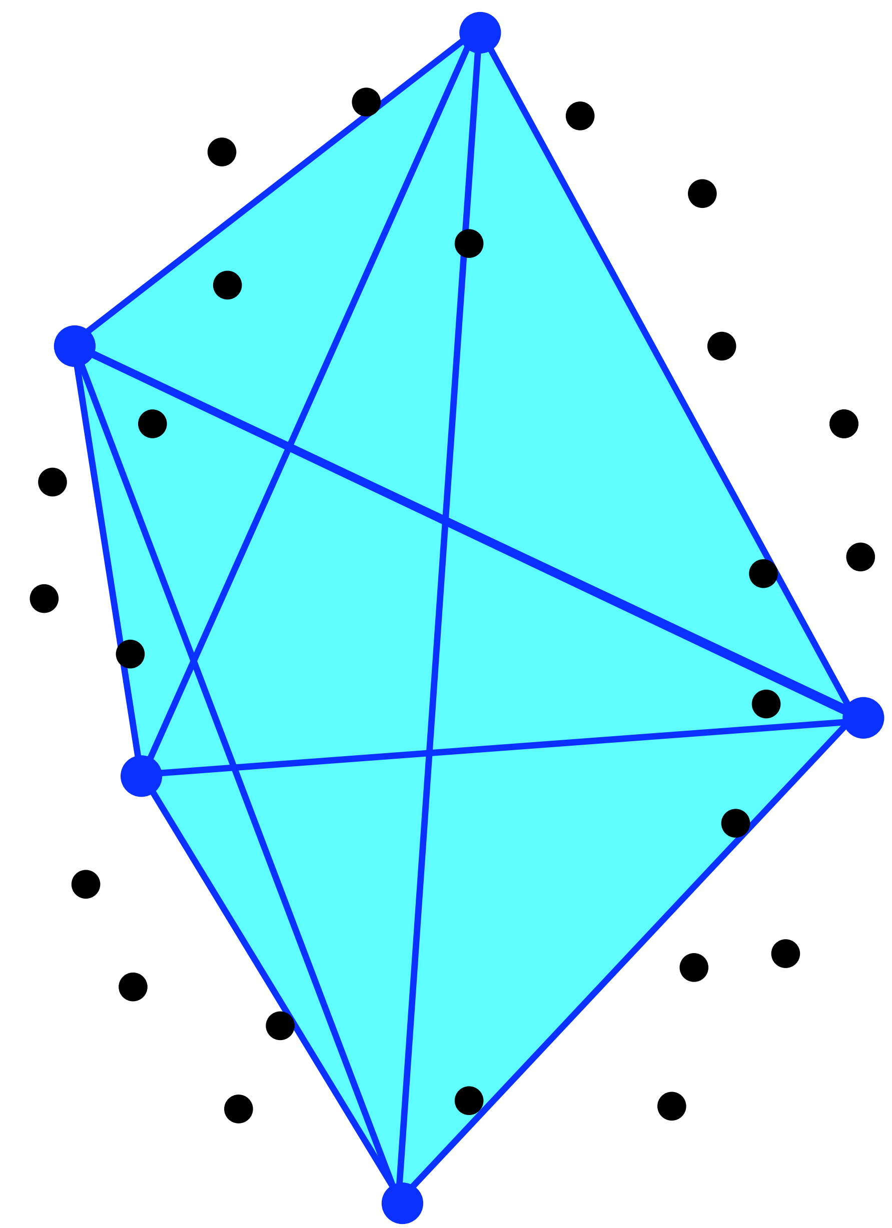
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w' \in W}$.
$\rightarrow$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

Init.: $L = \emptyset$; $\varepsilon = +\infty$

WHILE $L \subset W$
insert $p = argmax_{w \in W} d(w, L)$ in L
update $\varepsilon = \max_{w \in W} d(w, L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



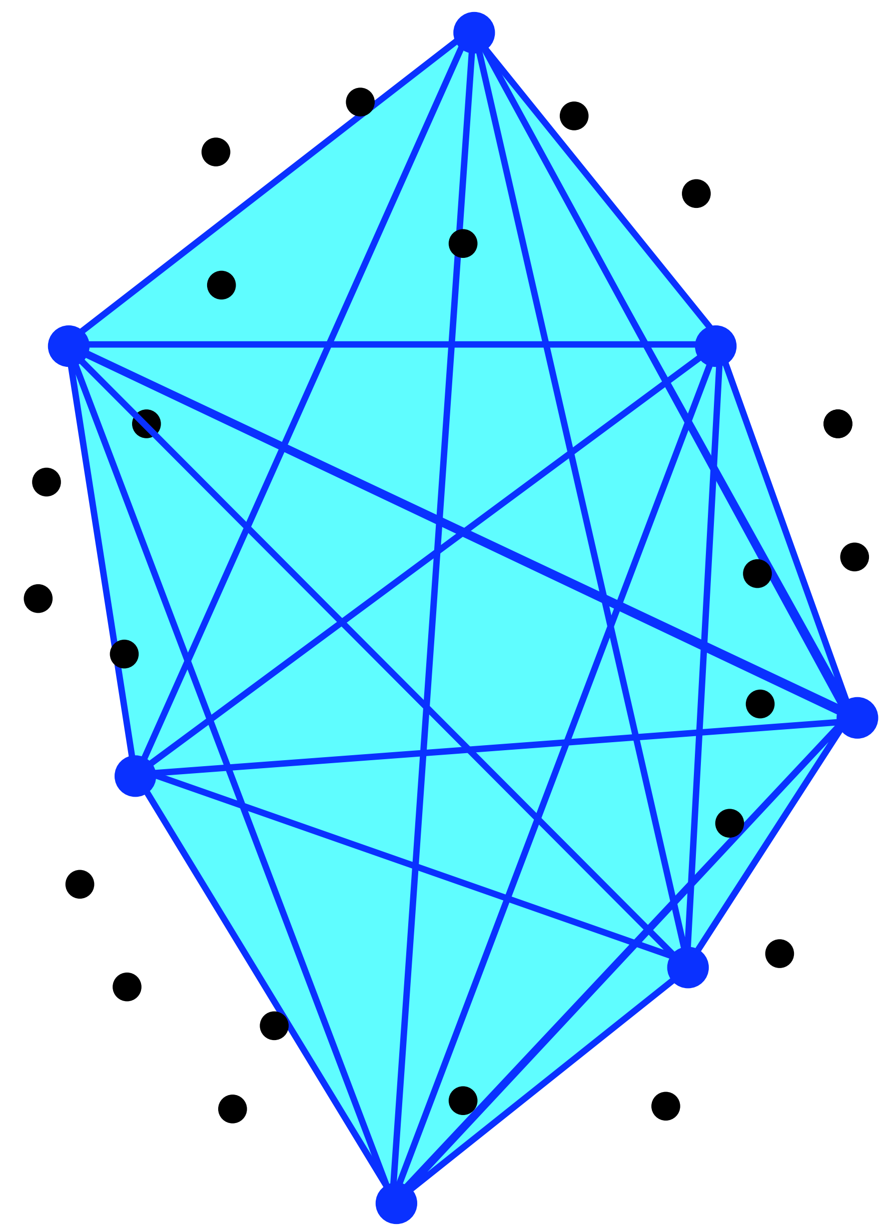
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w'\in W}$.
$\rightarrow$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

Init.: $L = \emptyset$; $\varepsilon = +\infty$

WHILE $L \subset W$
insert $p = argmax_{w\in W} d(w, L)$ in L
update $\varepsilon = \max_{w\in W} d(w, L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



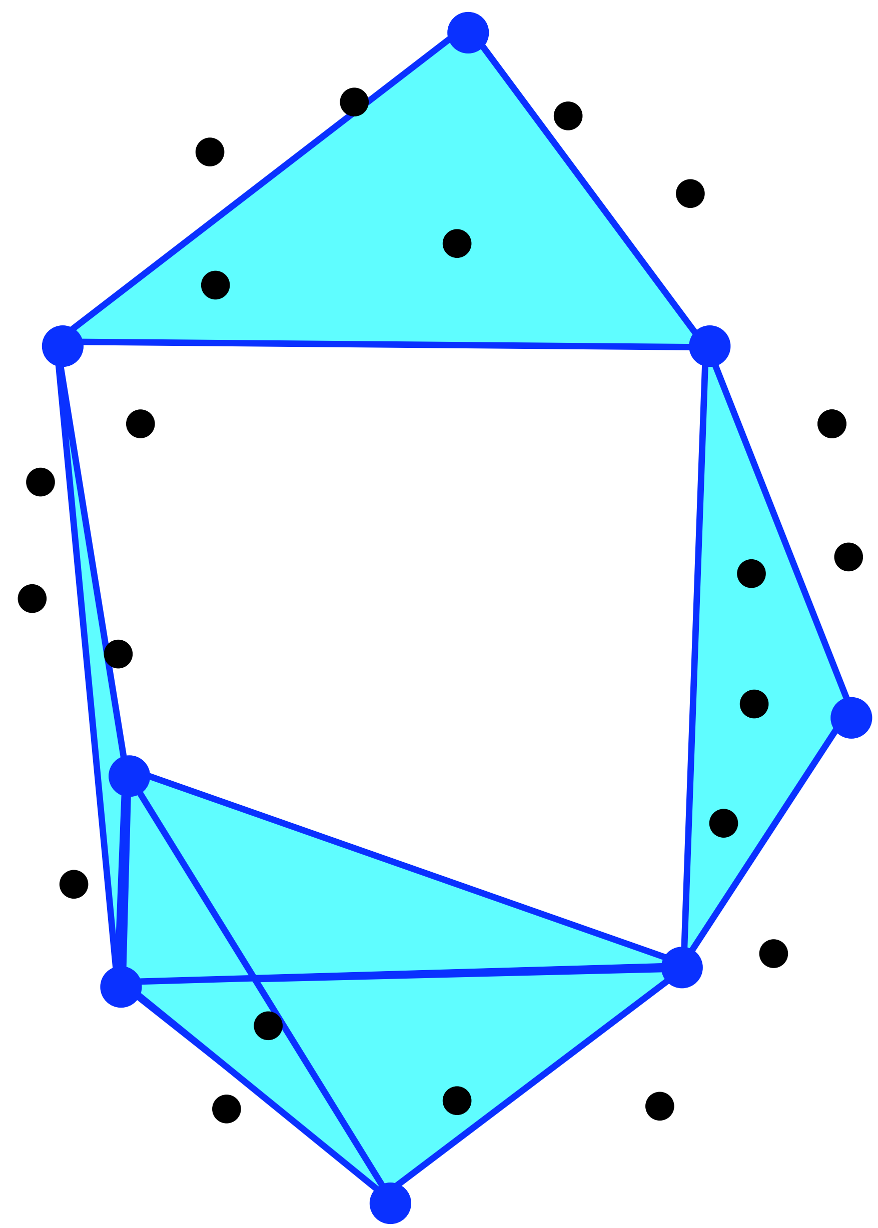
Multiscale inference

Input: A point cloud W and its pairewise distances $\{d(w,w')\}_{w,w'\in W}$.
$\rightarrow$ Maintain a nested pair $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$ where $L = L(\varepsilon)$.

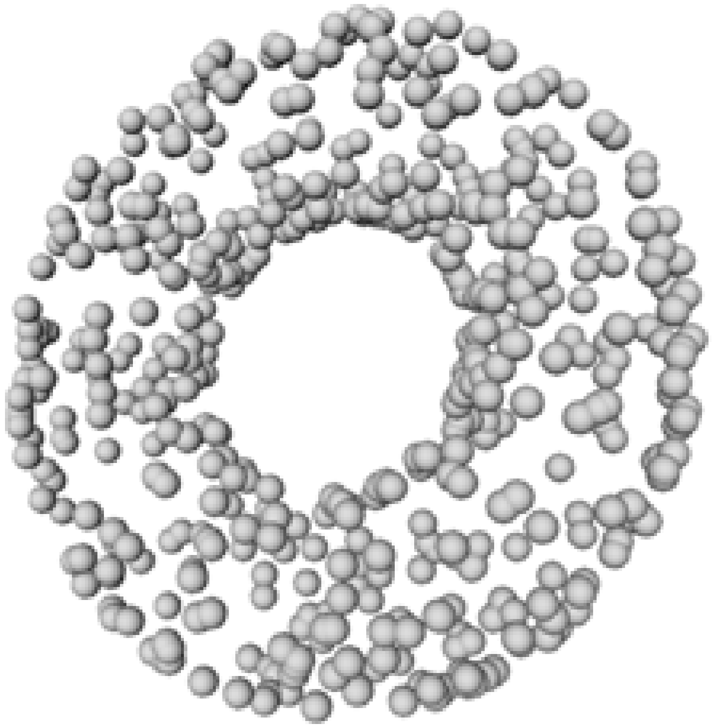
Init.: $L = \emptyset$; $\varepsilon = +\infty$

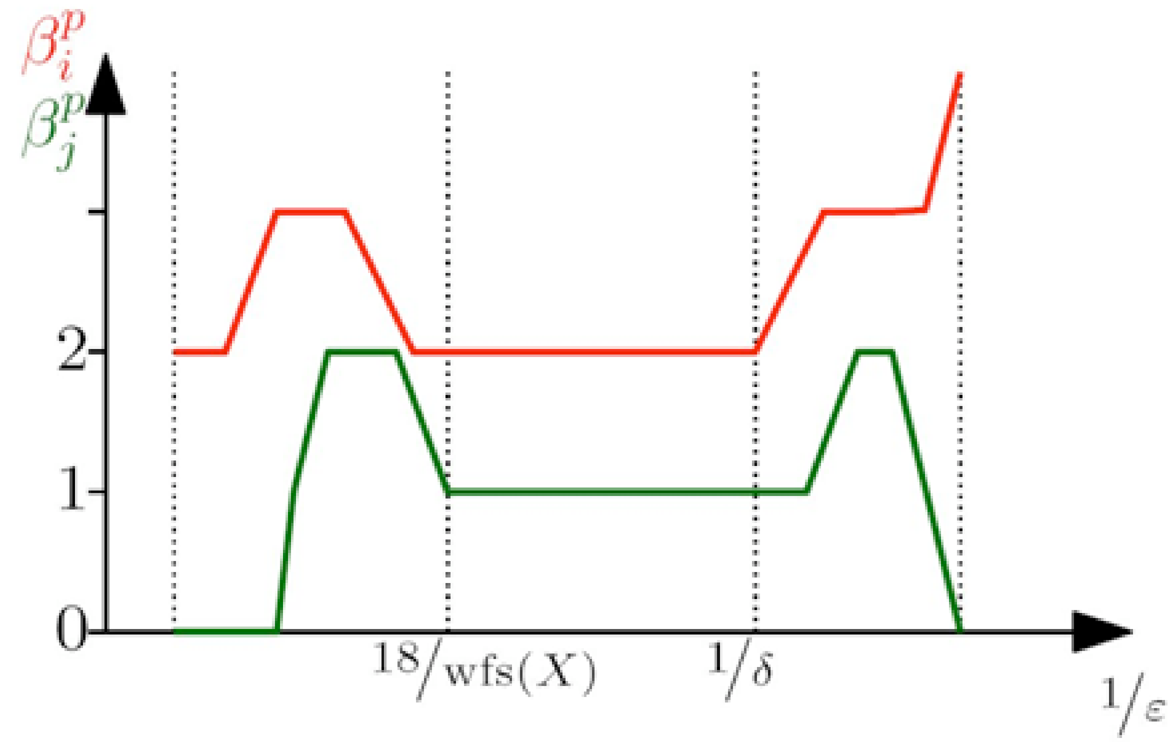
WHILE $L \subset W$
insert $p = argmax_{w\in W} d(w, L)$ in L
update $\varepsilon = \max_{w\in W} d(w, L)$
update $\mathcal{R}^{4\varepsilon}(L)$ and $\mathcal{R}^{16\varepsilon}(L)$
Persistence($\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$)
END_WHILE

Output: Sequence of persitent Betti numbers of $\mathcal{R}^{4\varepsilon}(L) \hookrightarrow \mathcal{R}^{16\varepsilon}(L)$



Multiscale inference



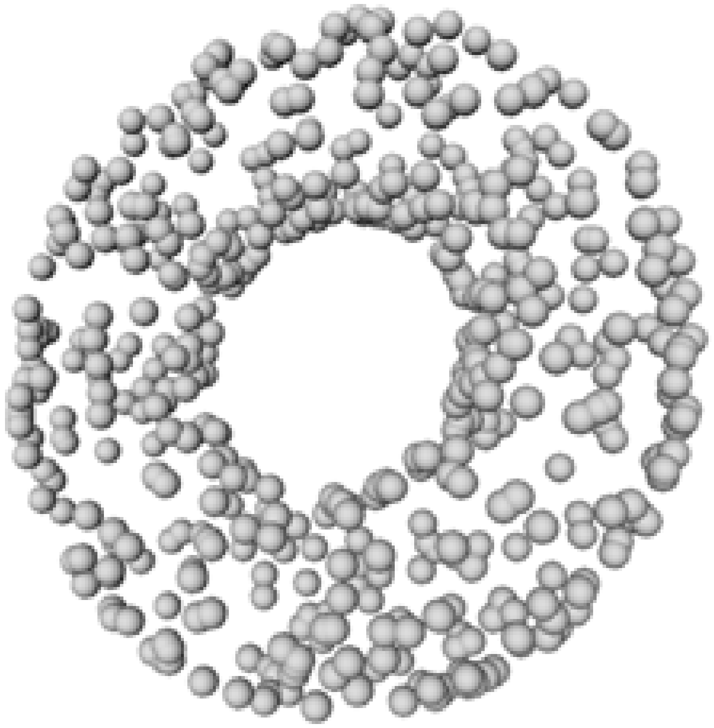


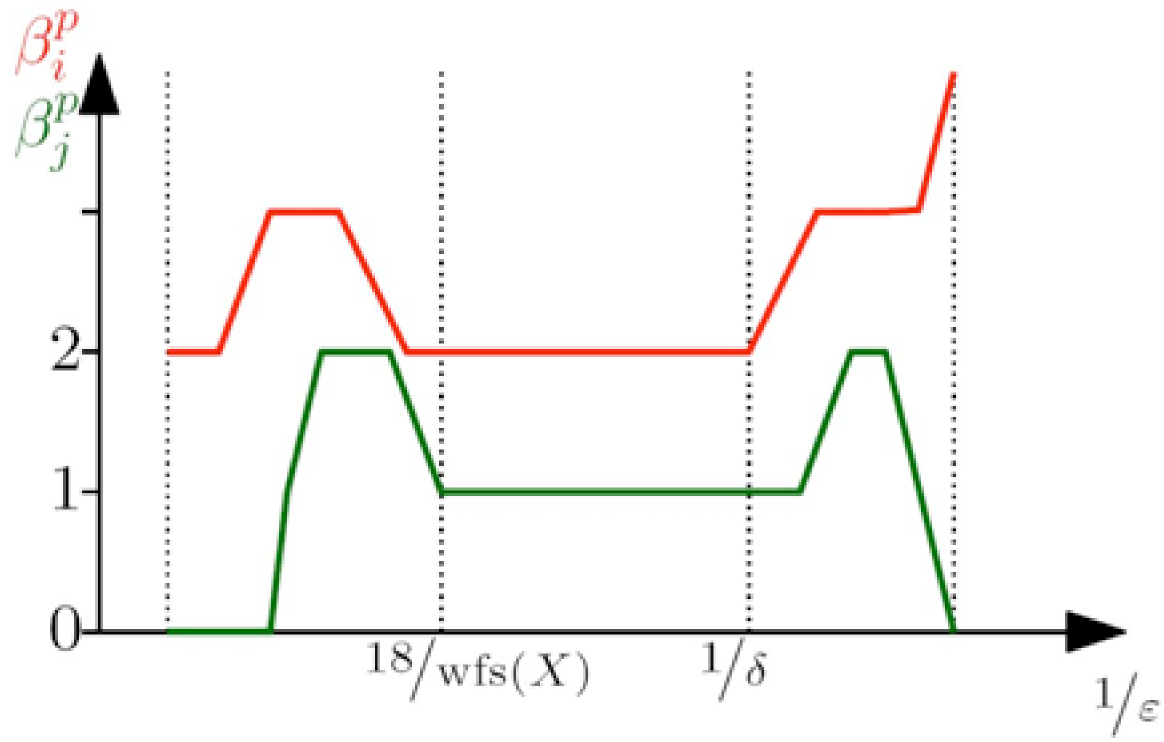
Theorem: [C-Oudot'08]
If $d_H(W, X) < \delta$ for $\delta < \frac{1}{18}\mathrm{wfs}(X)$, than at every iteration of the algorithm such that $\delta < \varepsilon < \frac{1}{18}\mathrm{wfs}(X)$,

$$\beta_k(X^\lambda) = \dim H_k(X^\lambda) = rk(H_k(\mathcal{R}^{4\varepsilon}(L)) \to H_k(\mathcal{R}^{4\varepsilon}(L)))$$

for any $\lambda \in (0, \mathrm{wfs}(X))$ and any $k \in \mathbb{N}$.

Multiscale inference





Complexity of the algorithm:

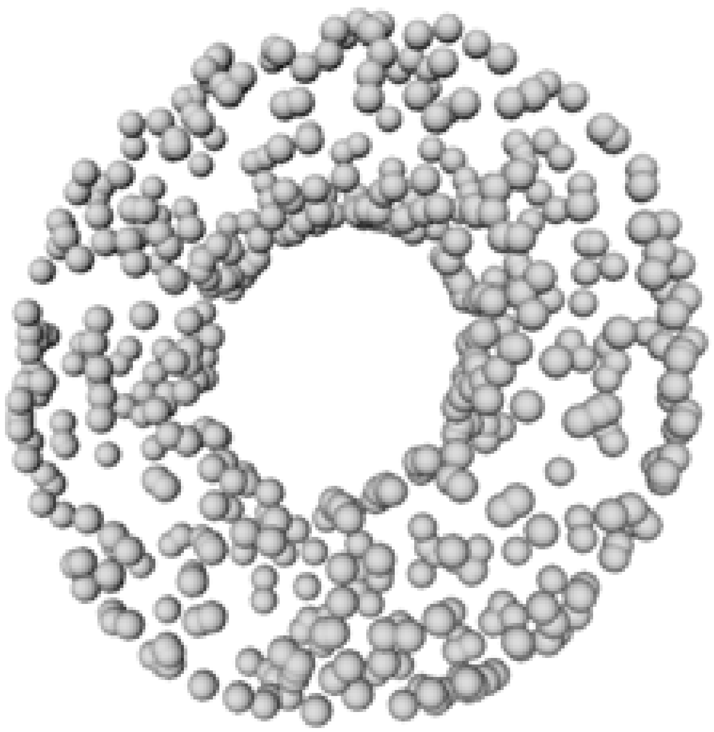
- If $X \subset \mathbb{R}^d$ is non smooth the running time of the algorithm is

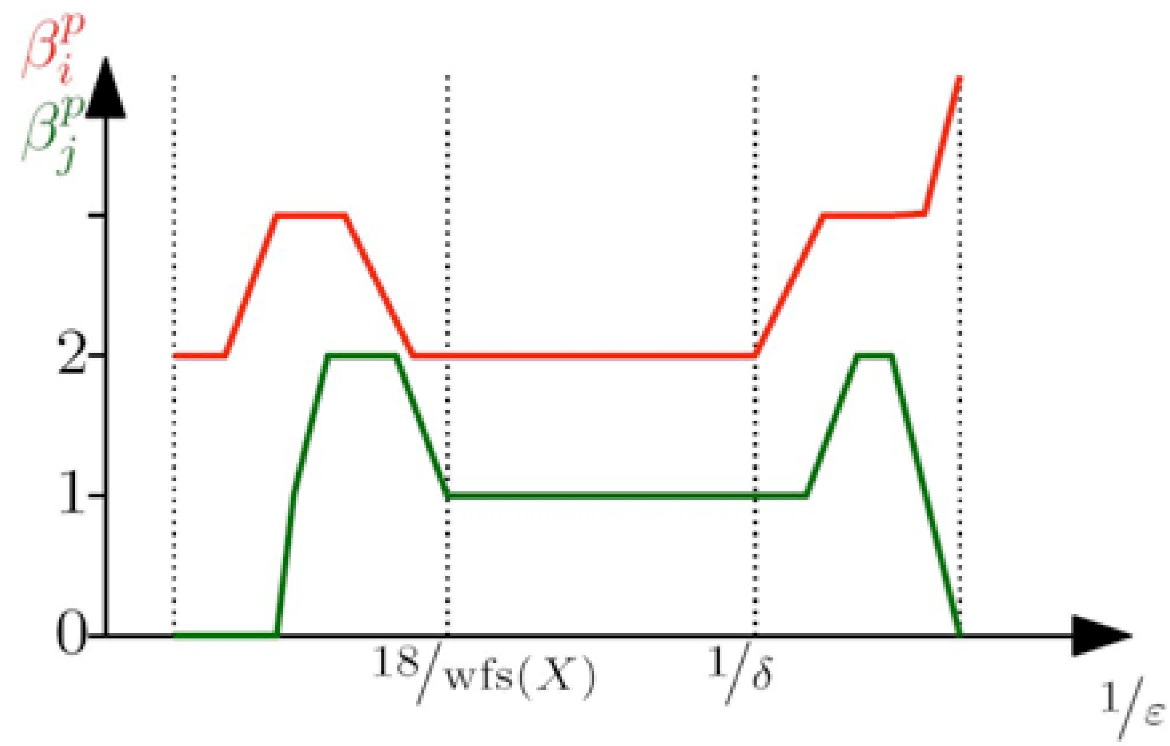
$$O(8^{33^d}|W|^5)$$

- If X is a smooth submanifold of $\mathbb{R}^d$ dimension m the running time is

$$O(8^{35^m}|W|)$$

Multiscale inference





Complexity of the algorithm:

- If $X \subset \mathbb{R}^d$ is non smooth the running time of the algorithm is

$$O(8^{33^d}|W|^5)$$

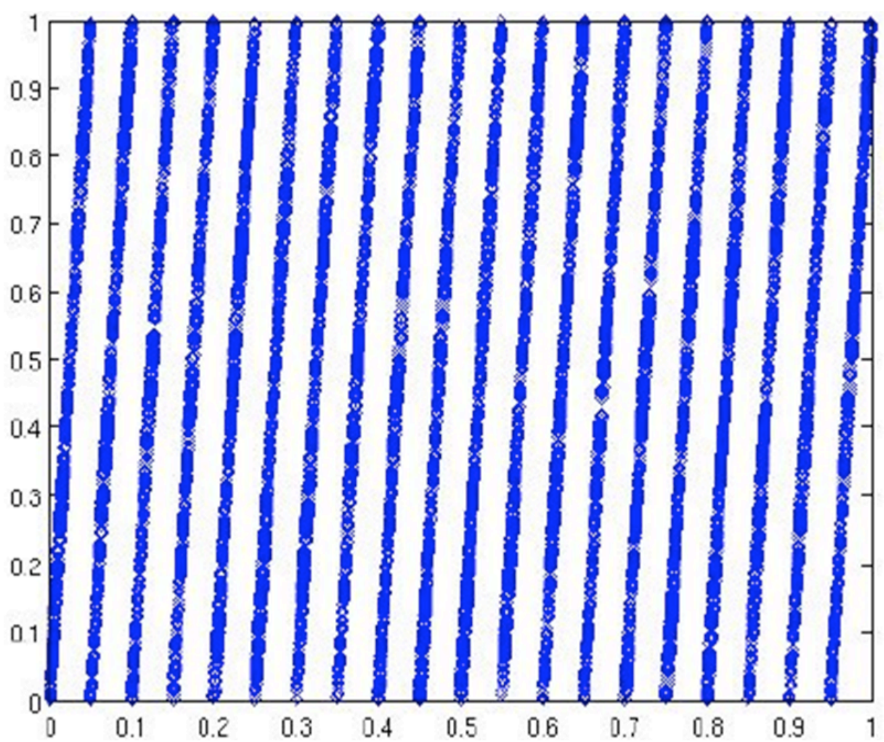
- If X is a smooth submanifold of $\mathbb{R}^d$ dimension m the running time is

$$O(8^{35^m}|W|)$$

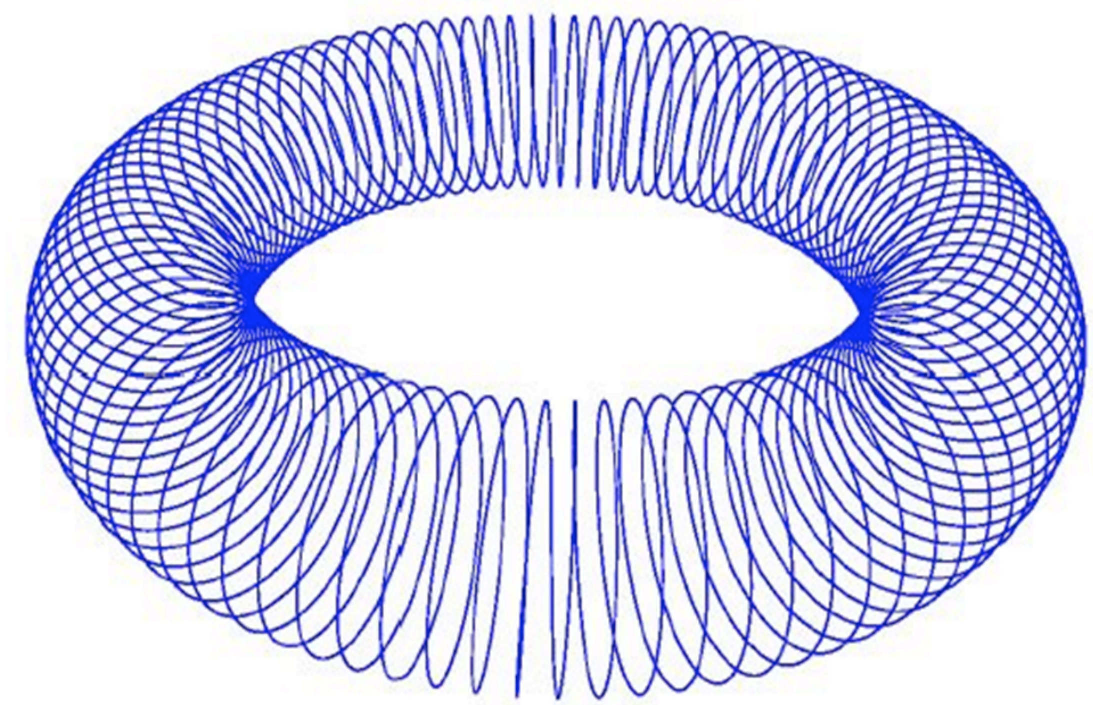
Depend on the intrinsic dimension of X

A synthetic example

$[0, 1] \times [0, 1]$



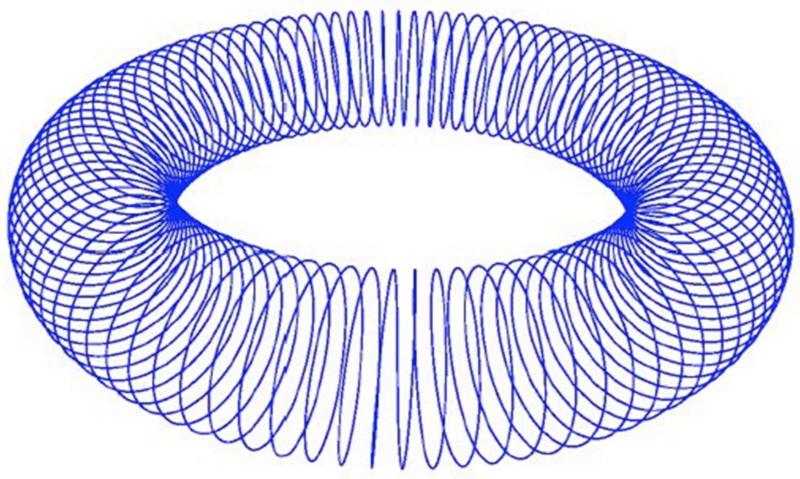
$\mathbb{R}^{1000}$

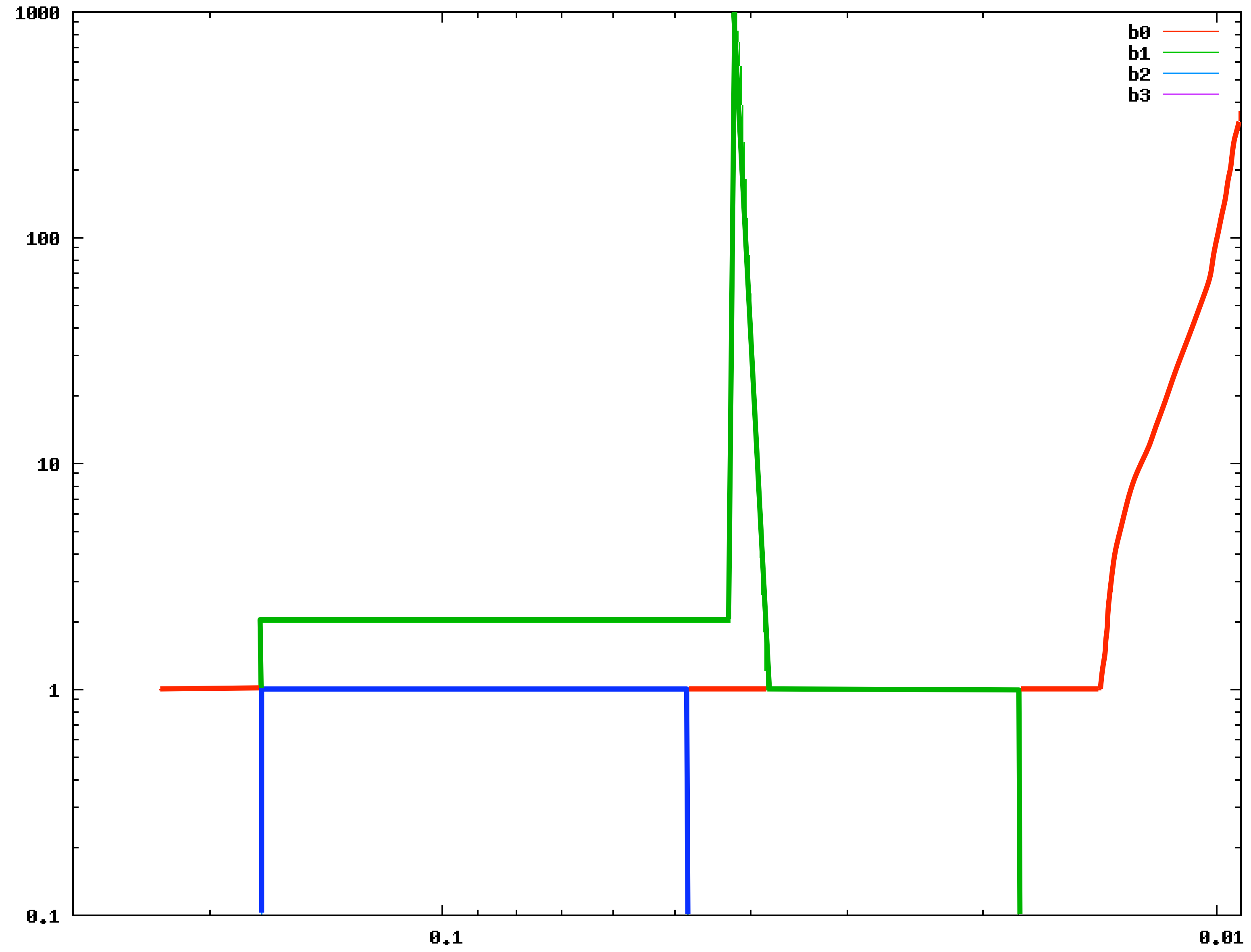


Non-linear embedding of $\mathbb{S}^1 \times \mathbb{S}^1$ in $\mathbb{R}^{1000}$

$50,000$ points sampled uniformly at random from a curve drawn on the 2-torus $\mathbb{S}^1 \times \mathbb{S}^1$.

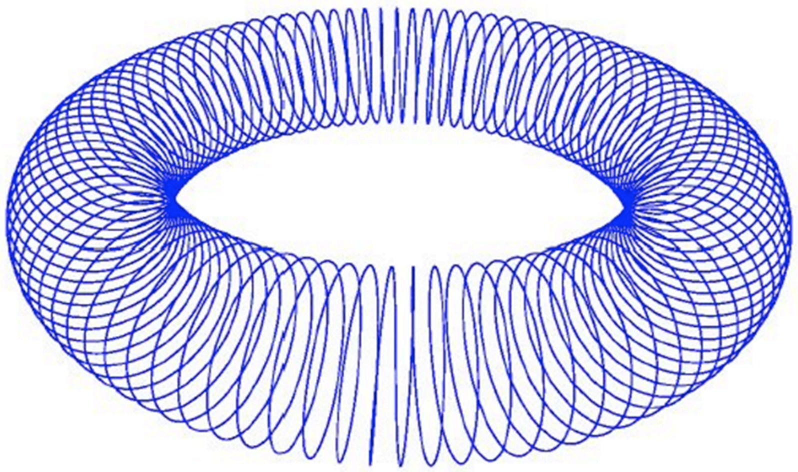
A synthetic example

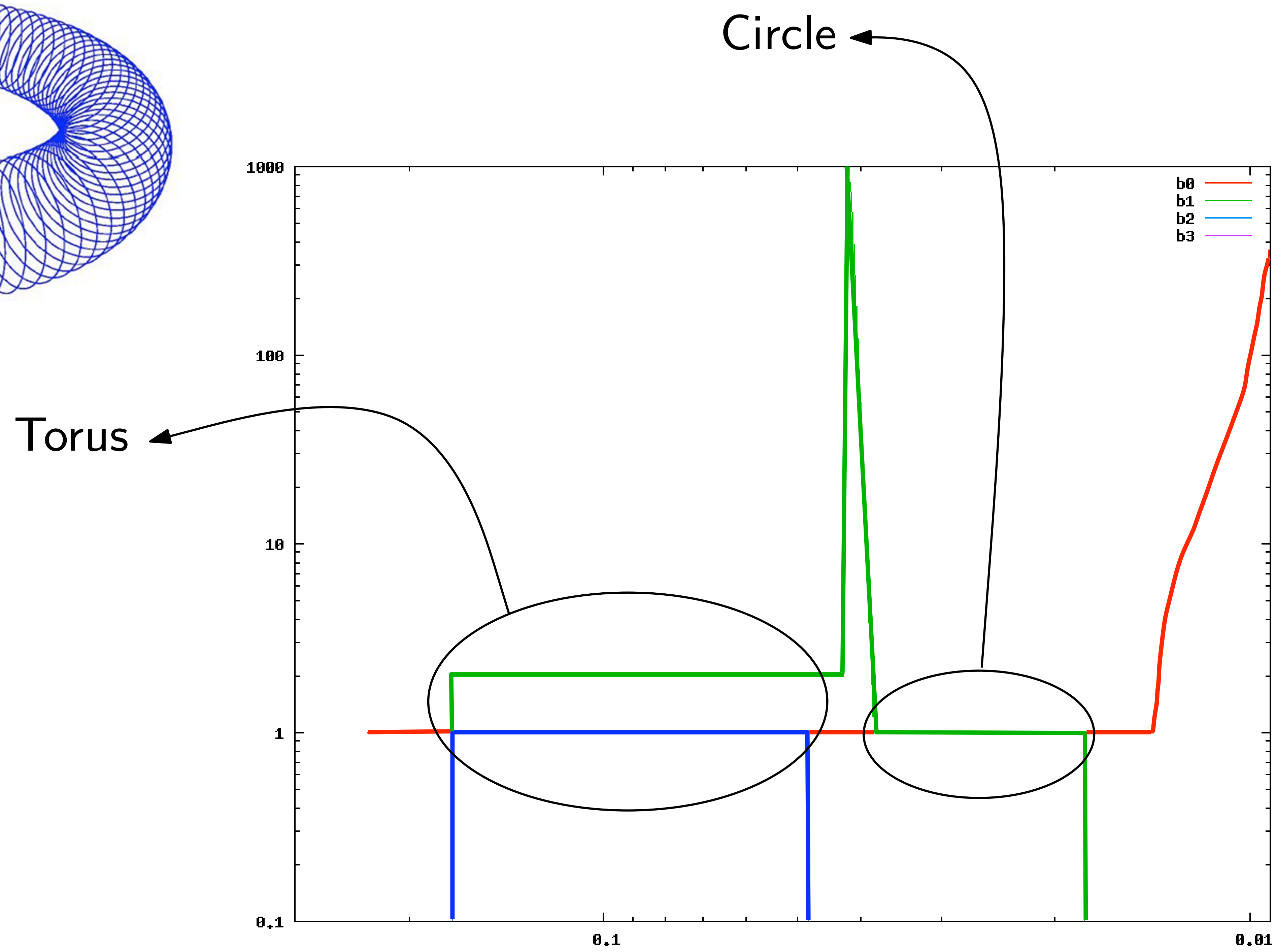




Output: sequence of Betti numbers on a log-log scale

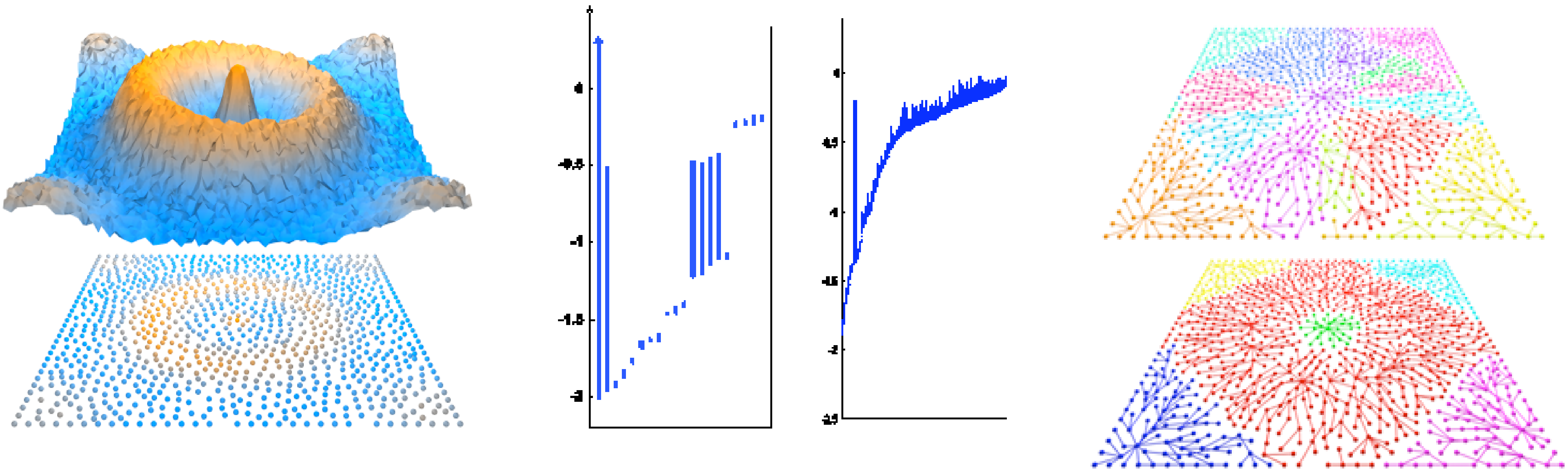
A synthetic example





Output: sequence of Betti numbers on a log-log scale

Generalization(s)

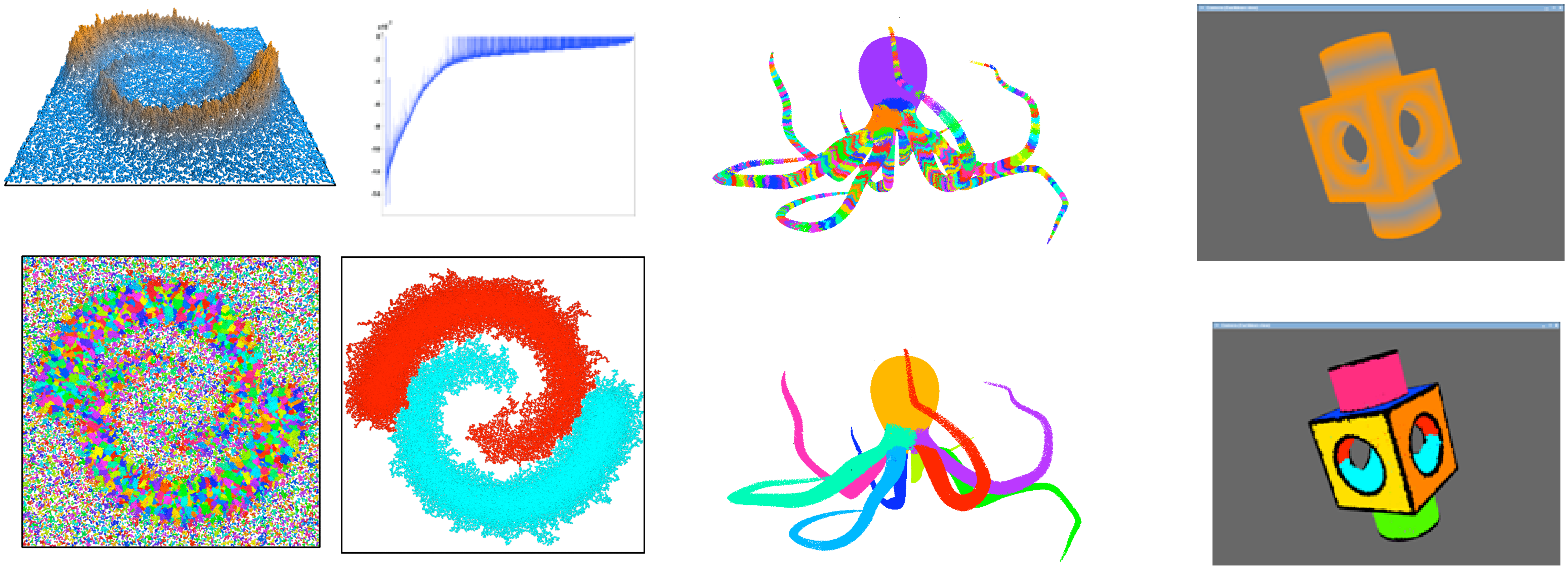


Previous approach can be generalized, leading to robust algorithms to compute the topological persistence of functions defined over point clouds sampled around unknown shapes

Ref:

- F. Chazal, L. Guibas, S. Oudot, P. Skraba, *Analysis of Scalar Fields over Point Cloud Data*, proc. ACM Symposium on Discrete Algorithms 2009.
- F. Chazal, S. Oudot, *Toward Persistence-Based Reconstruction in Euclidean Spaces*, proc. ACM Symposium on Computational Geometry 2008.

Generalization(s)



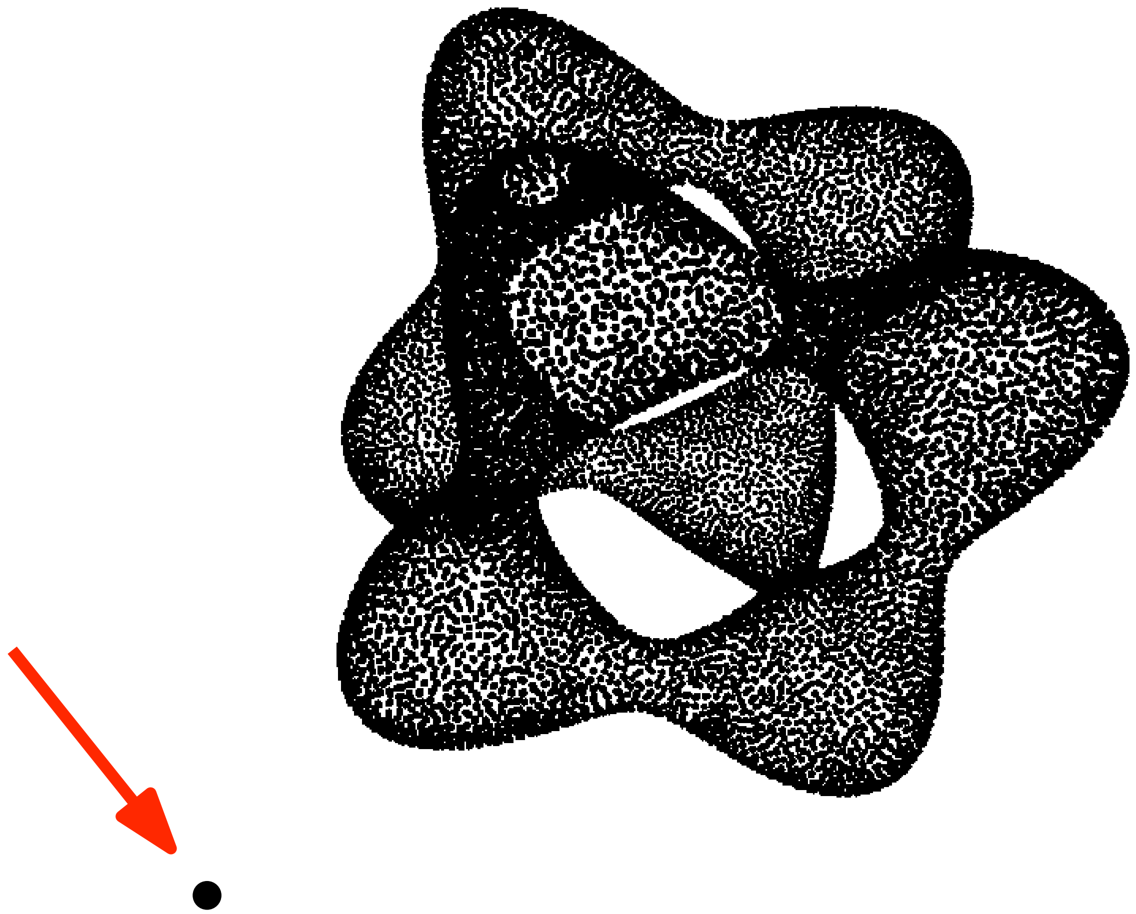
Applications to clustering, segmentations, sensor networks,...

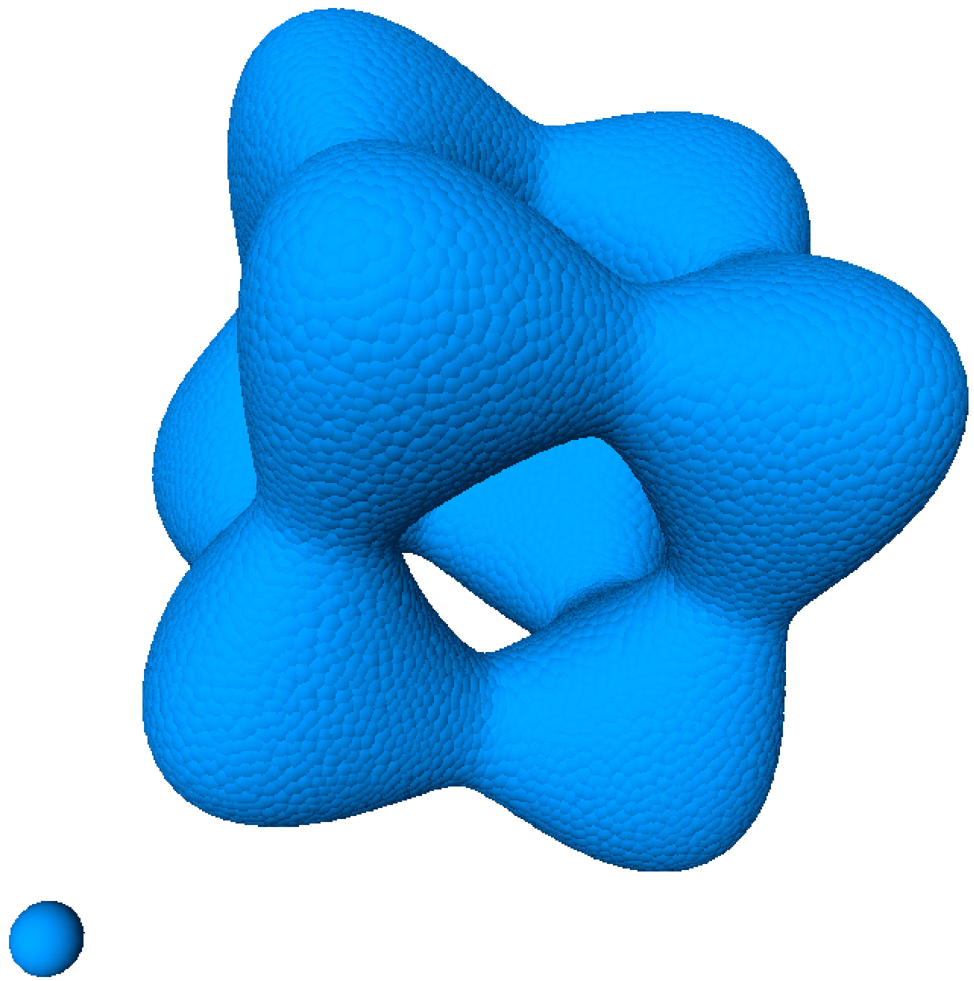
Ref:

- F. Chazal, L. Guibas, S. Oudot, P. Skraba, *Analysis of Scalar Fields over Point Cloud Data*, proc. ACM Symposium on Discrete Algorithms 2009.

- F. Chazal, S. Oudot, *Toward Persistence-Based Reconstruction in Euclidean Spaces*, proc. ACM Symposium on Computational Geometry 2008.

The problem of “outliers”





If $K' = K \cup \{x\}$ where $d_K(x) > R$, then $\|d_K - d_{K'}\|_\infty > R$: offset-based inference methods fail!

Question: Can we generalized the previous approach by replacing the distance function by a “distance-like” function having a better behavior with respect to “'noise” and “outliers”?

Distance-like functions: the three main ingredients of stability

- the stability of the map $K \mapsto d_K$:
  $\|d_K - d_{K'}\|_\infty = d_H(K, K')$

Distance-like functions: the three main ingredients of stability

- the stability of the map $K \mapsto d_K$:
  $\|d_K - d_{K'}\|_\infty = d_H(K, K')$

- the 1-Lipschitz property for d_K; ⟶ d_K is differentiable almost everywhere.

Distance-like functions: the three main ingredients of stability

- the stability of the map $K \mapsto d_K$:
  $\|d_K - d_{K'}\|_\infty = d_H(K, K')$

- the 1-Lipschitz property for d_K; $\longrightarrow$ d_K is differentiable almost everywhere.

- the 1-concavity of the function d_K^2:
  $x \to \|x\|^2 - d_K^2(x)$ is convex. $\longrightarrow$
  - the gradient vector field ∇d_K is well defined and integrable (although not continuous).
  - d_K admits a second derivative almost everywhere.

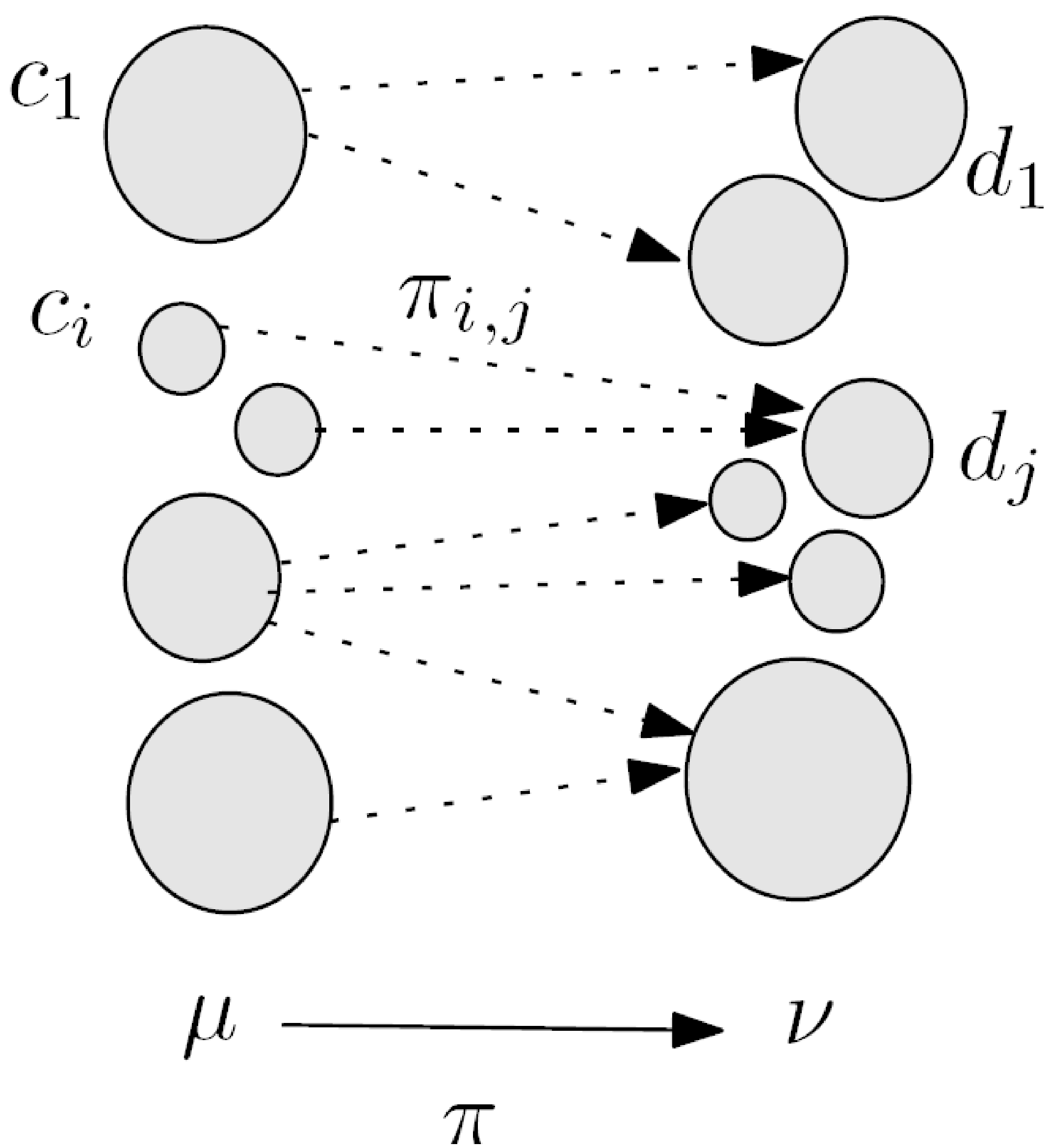
Replacing compact sets by measures

A **measure** μ is a mass distribution on $\mathbb{R}^d$:
mathematically, it is defined as a map μ that takes a (Borel) subset $B \subset \mathbb{R}^d$ and outputs a nonnegative number $\mu(B)$. Moreover we ask that if (B_i) are disjoint subsets, $\mu\left(\bigcup_{i\in\mathbb{N}} B_i\right) = \sum_{i\in\mathbb{N}} \mu(B_i)$.

- $\mu(B)$ corresponds to to the mass of μ contained in B
- a point cloud $C = \{p_1, \ldots, p_n\}$ defines a measure $\mu_C = \frac{1}{n}\sum_i \delta_{p_i}$
- the volume form on a k-dimensional submanifold M of $\mathbb{R}^d$ defines a measure $\mathrm{vol}_{k|M}$.
- etc...

Distance between measures

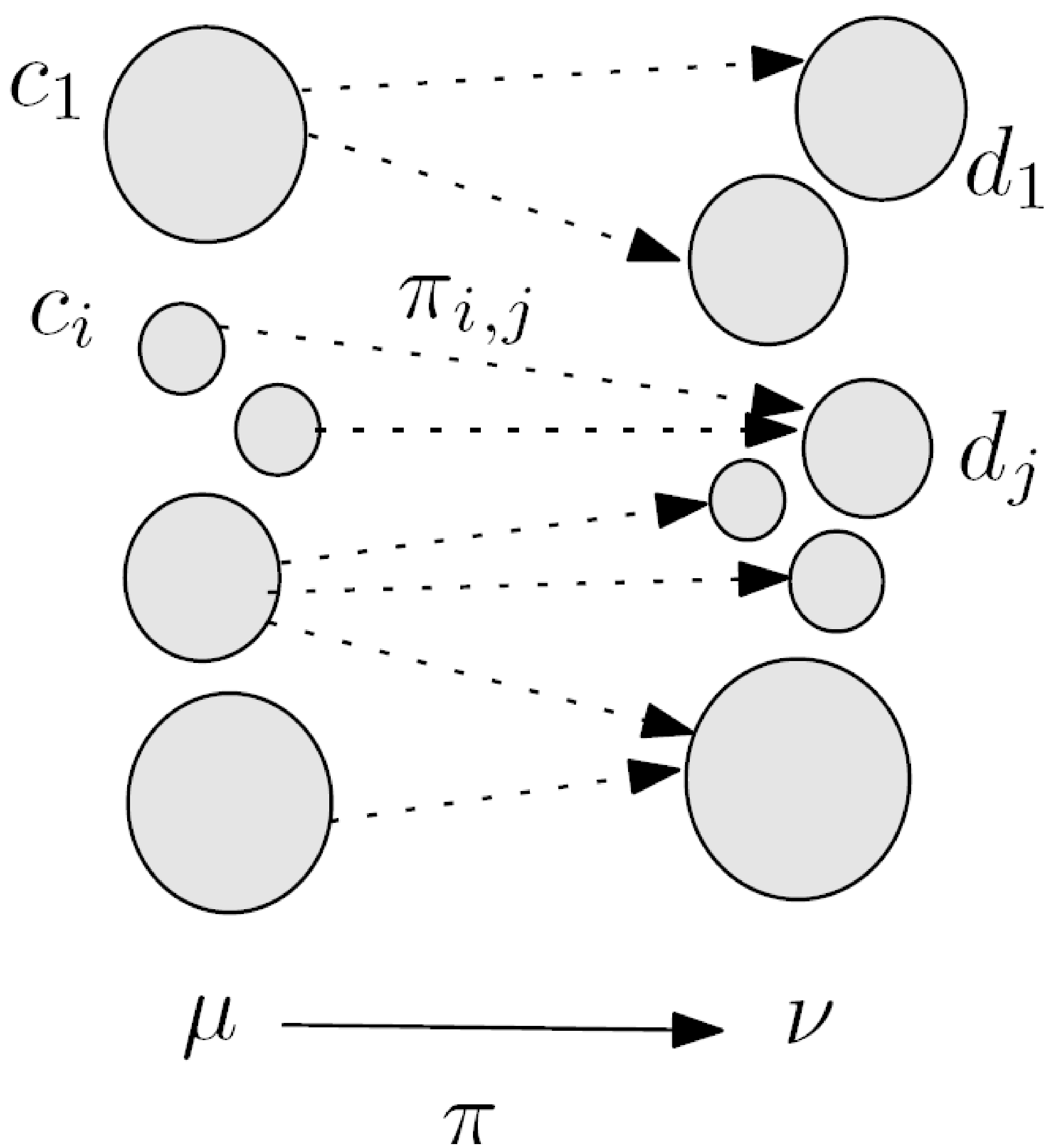
The **Wasserstein distance** $d_W(\mu,\nu)$ between two probability measures μ,ν quantifies the optimal cost of pushing μ onto ν, the cost of moving a small mass dx from x to y being $\|x-y\|^2 dx$.



1. μ and ν are discrete measures: $\mu = \sum_i c_i \delta_{x_i}$, $\nu = \sum_j d_j \delta_{y_j}$ with $\sum_j d_j = \sum_i c_i$.
2. *Transport plan*: set of coefficients $\pi_{ij} \geq 0$ with $\sum_i \pi_{ij} = d_j$ and $\sum_j \pi_{ij} = c_i$.
3. Cost of a transport plan
   $$C(\pi) = \left(\sum_{ij} \|x_i - y_j\|^2 \pi_{ij} \right)^{1/2}$$
4. $d_W(\mu,\nu) := \inf_\pi C(\pi)$

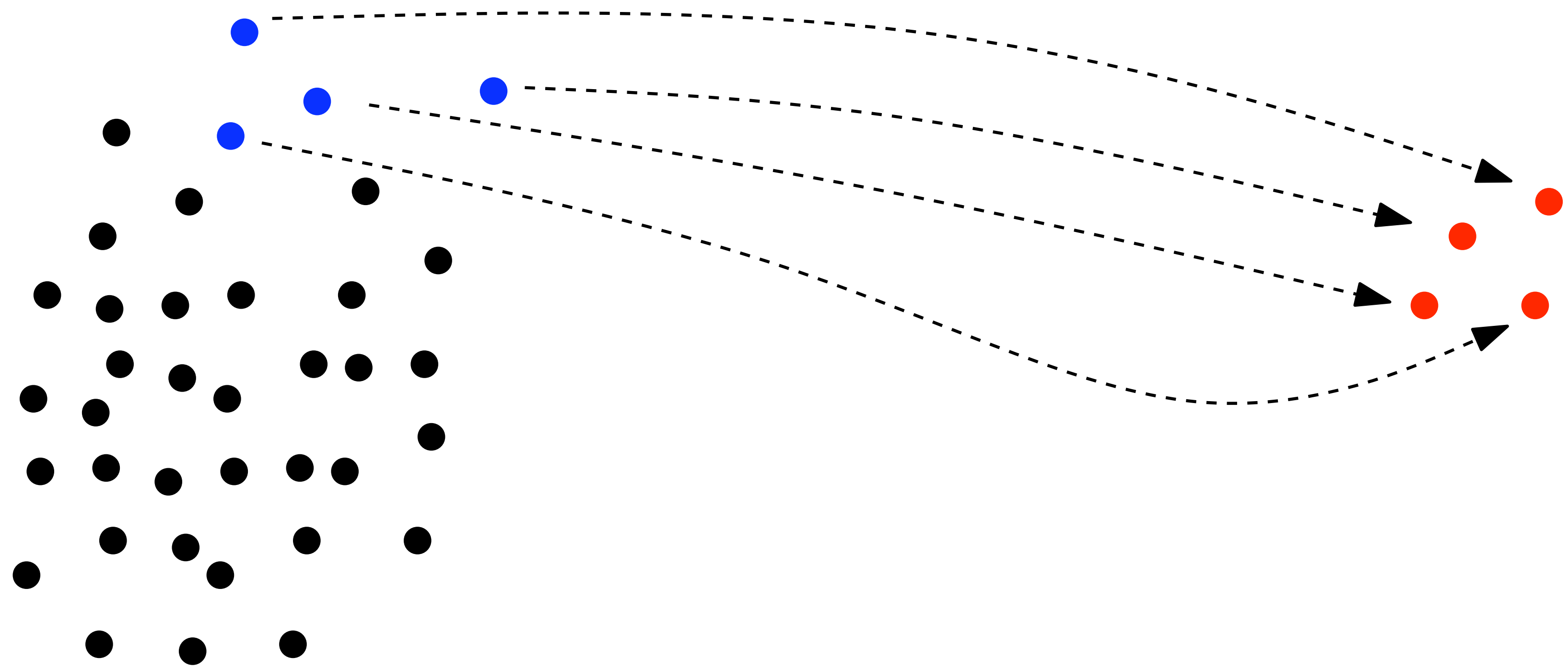
Distance between measures

The **Wasserstein distance** $d_W(\mu, \nu)$ between two probability measures μ, ν quantifies the optimal cost of pushing μ onto ν, the cost of moving a small mass dx from x to y being $\|x - y\|^2 dx$.



1. μ and ν are proba measures in $\mathbb{R}^d$
2. *Transport plan*: π a proba measure on $\mathbb{R}^d \times \mathbb{R}^d$ s.t. $\pi(A \times \mathbb{R}^d) = \mu(A)$ and $\pi(\mathbb{R}^d \times B) = \nu(B)$.
3. Cost of a transport plan
   $C(\pi) = \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^2 d\pi(x, y) \right)^{\frac{1}{2}}$
4. $d_W(\mu, \nu) := \inf_\pi C(\pi)$

Wasserstein distance



Examples:

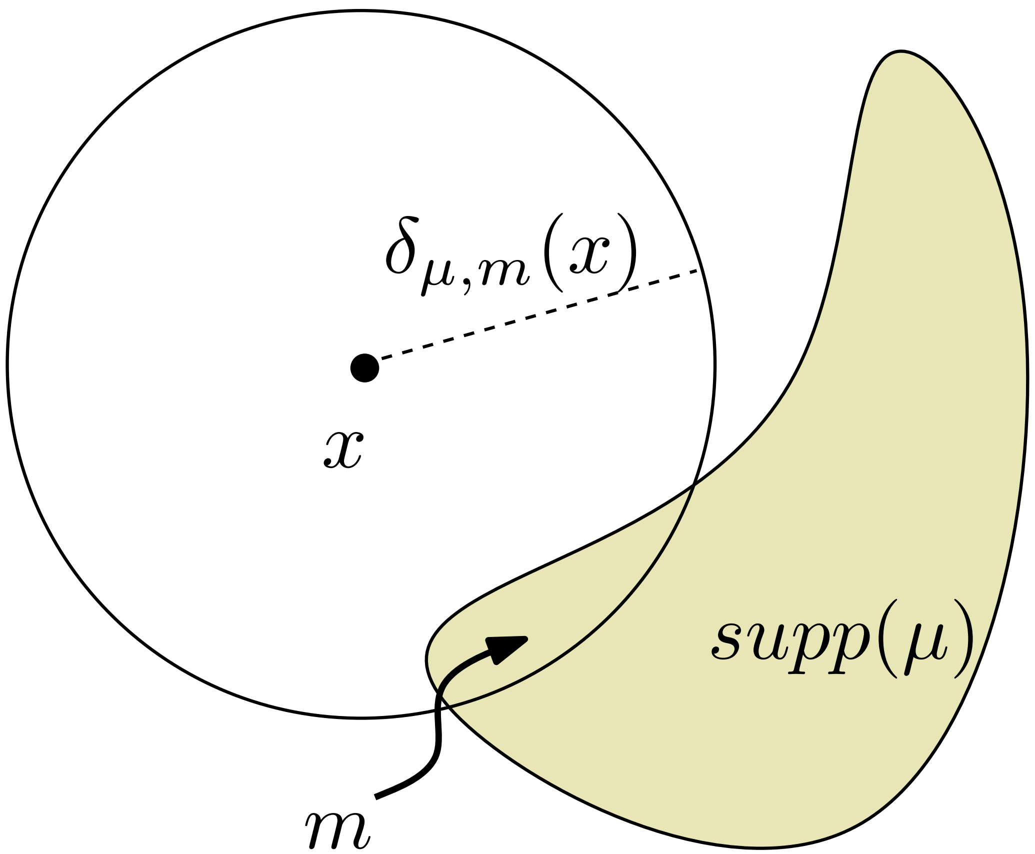
- If C_1 and C_2 are two point clouds, with $\#C_1 = \#C_2$, then $d_W(\mu_{C_1}, \mu_{C_2})$ is the square root of the cost of a minimal least-square matching between C_1 and C_2.

- If $C = \{p_1, \ldots, p_n\}$ is a point cloud, and $C' = \{p_1, \ldots, p_{n-k-1}, o_1, \ldots, o_k\}$ with $d(o_i, C) = R$, then

$$d_H(C, C') \geq R \qquad \text{but} \qquad d_W(\mu_C, \mu_{C'}) \leq \frac{k}{n}(R + \mathrm{diam}(C))$$

The distance to a measure

Distance function to a measure, first attempt:
Let $m \in]0,1[$ be a positive mass, and μ a probability measure on $\mathbb{R}^d$: $\delta_{\mu,m}(x) = \inf\{r > 0 : \mu(\mathbb{B}(x,r)) > m\}$.



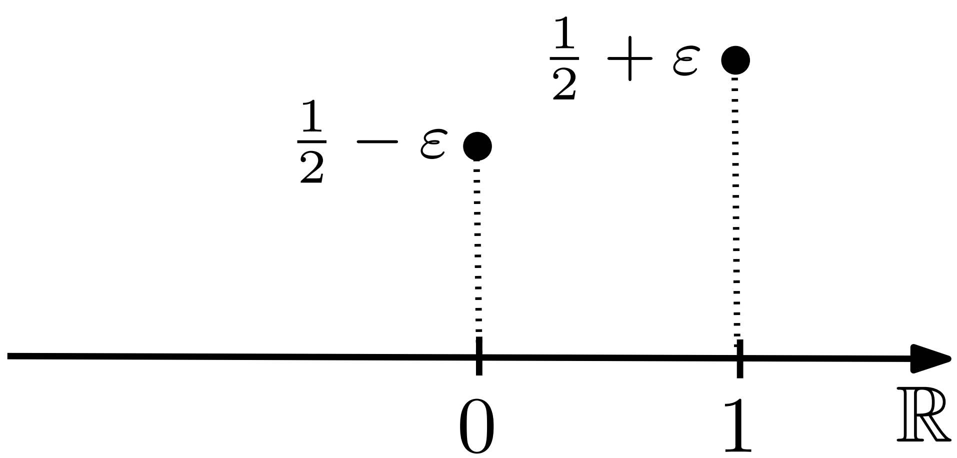
- $\delta_{\mu,m}$ is the smallest distance needed to capture a mass of at least m;
- Coincides with the distance to the k-th neighbor when $m = k/n$ and $\mu = \frac{1}{n}\sum_{i=1}^{n} \delta_{p_i}$:

$$\delta_{\mu,k/n}(\mu) = \|x - p_C^k(x)\|$$

Unstability of $\mu \mapsto \delta_{\mu,m}$

Distance function to a measure, first attempt:
Let $m \in]0,1[$ be a positive mass, and μ a probability measure on $\mathbb{R}^d$: $\delta_{\mu,m}(x) = \inf\{r > 0 : \mu(\mathbb{B}(x,r)) > m\}$.

Unstability under Wasserstein perturbations:



$$\mu_\varepsilon = (1/2 - \varepsilon)\delta_0 + (1/2 + \varepsilon)\delta_1$$

for $\varepsilon > 0 :\ \forall x < 0,\ \delta_{\mu_\varepsilon,1/2}(x) = |x - 1|$

for $\varepsilon = 0 :\ \forall x < 0,\ \delta_{\mu_0,1/2}(x) = |x - 0|$

Consequence: the map $\mu \mapsto \delta_{\mu,m} \in \mathcal{C}^0(\mathbb{R}^d)$ is discontinuous whatever the (reasonable) topology on $\mathcal{C}^0(\mathbb{R}^d)$.

The distance function to a measure

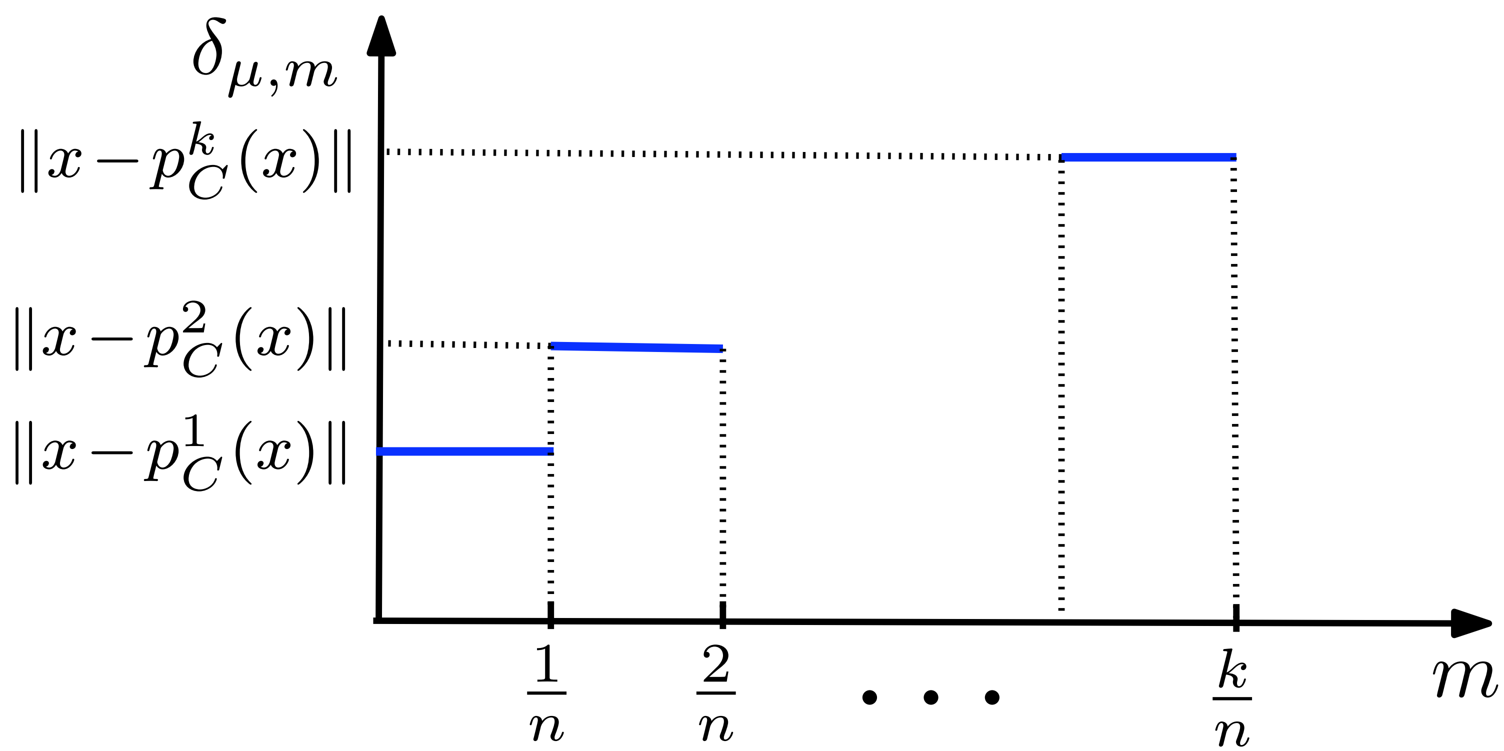
Definition: If μ is a probability measure on $\mathbb{R}^d$ and $m_0 > 0$, one let:

$$d_{\mu,m_0} : x \in \mathbb{R}^d \mapsto \left(\frac{1}{m_0} \int_0^{m_0} \delta_{\mu,m}^2(x) dm \right)^{1/2}$$

The distance function to a measure

Definition: If μ is a probability measure on $\mathbb{R}^d$ and $m_0 > 0$, one let:

$$d_{\mu,m_0} : x \in \mathbb{R}^d \mapsto \left(\frac{1}{m_0} \int_0^{m_0} \delta^2_{\mu,m}(x) dm \right)^{1/2}$$



Example. Let $C = \{p_1, \dots, p_n\}$ and $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{p_i}$. Let $p_C^k(x)$ denote the kth nearest neighbor to x in C, and set $m_0 = k_0/n$:

$$d_{\mu,m_0}(x) = \left(\frac{1}{k_0} \sum_{k=1}^{k_0} \|x - p_C^k(x)\|^2 \right)^{1/2}$$

Another expression for d_{μ,m_0}

$$d_{\mu,m_0}(x) = \min_{\tilde{\mu}} \left\{ d_W\left(\delta_x, \frac{1}{m_0}\tilde{\mu}\right) : \tilde{\mu}(\mathbb{R}^d) = m_0 \text{ and } \tilde{\mu} \leq \mu \right\}$$

"The projection submeasure": $\tilde{\mu}_{x,m_0}$ = the restriction of μ on the ball $B = \mathbb{B}(x, \delta_{\mu,m_0}(x))$, whose trace on the sphere ∂B has been rescaled so that the total mass of $\tilde{\mu}_{x,m_0}$ is m_0.

$$d^2_{\mu,m_0}(x) = \frac{1}{m_0} \int_{h \in \mathbb{R}^d} \|h - x\|^2 \, d\tilde{\mu}_{x,m_0} = d^2_W\left(\delta_x, \frac{1}{m_0}\tilde{\mu}_{x,m_0}\right)$$

Another expression for d_{μ,m_0}

$$d_{\mu,m_0}(x) = \min_{\tilde{\mu}} \left\{ d_W\left(\delta_x, \frac{1}{m_0}\tilde{\mu}\right) : \tilde{\mu}(\mathbb{R}^d) = m_0 \text{ and } \tilde{\mu} \leq \mu \right\}$$

Proof:

Another expression for d_{μ,m_0}

$$d_{\mu,m_0}(x) = \min_{\tilde{\mu}} \left\{ d_W\left(\delta_x, \frac{1}{m_0}\tilde{\mu}\right) : \tilde{\mu}(\mathbb{R}^d) = m_0 \text{ and } \tilde{\mu} \leq \mu \right\}$$

Proof:

Only one transport plan : $y \in \mathbb{R}^d \to x$

$$\int_{\mathbb{R}^d} \|h - x\|^2 d\tilde{\mu}(h)$$

Another expression for d_{μ,m_0}

$$d_{\mu,m_0}(x) = \min_{\tilde{\mu}} \left\{ d_W\left(\delta_x, \frac{1}{m_0}\tilde{\mu}\right) : \tilde{\mu}(\mathbb{R}^d) = m_0 \text{ and } \tilde{\mu} \leq \mu \right\}$$

Proof:

$$\int_{\mathbb{R}^d} \|h-x\|^2 d\tilde{\mu}(h) = \int_{\mathbb{R}_+} t^2 d\tilde{\mu}_x(t) = \int_0^{m_0} F_{\tilde{\mu}_x}^{-1}(m)^2 dm$$

pushforward of $\tilde{\mu}$ by the distance function to x.

$F_{\tilde{\mu}_x}(t) = \tilde{\mu}_x([0,t))$ is the cumulative function of $\tilde{\mu}_x$ and $F_{\tilde{\mu}_x}^{-1}(m) = \inf\{t \in \mathbb{R} : F_{\tilde{\mu}_x}(t) > m\}$ is its generalized inverse

Another expression for d_{μ,m_0}

$$d_{\mu,m_0}(x) = \min_{\tilde{\mu}} \left\{ d_W\left(\delta_x, \frac{1}{m_0}\tilde{\mu}\right) : \tilde{\mu}(\mathbb{R}^d) = m_0 \text{ and } \tilde{\mu} \leq \mu \right\}$$

Proof:

$$\int_{\mathbb{R}^d} \|h - x\|^2 d\tilde{\mu}(h) = \int_{\mathbb{R}_+} t^2 d\tilde{\mu}_x(t) = \int_0^{m_0} F_{\tilde{\mu}_x}^{-1}(m)^2 dm$$

pushforward of $\tilde{\mu}$ by the distance function to x.

$F_{\tilde{\mu}_x}(t) = \tilde{\mu}_x([0,t))$ is the cumulative function of $\tilde{\mu}_x$ and $F_{\tilde{\mu}_x}^{-1}(m) = \inf\{t \in \mathbb{R} : F_{\tilde{\mu}_x}(t) > m\}$ is its generalized inverse

- $\tilde{\mu} \leq \mu \Rightarrow F_{\tilde{\mu}_x}(t) \leq F_{\mu_x}(t) \Rightarrow F_{\tilde{\mu}_x}^{-1}(m) \geq F_{\mu_x}^{-1}(m)$
- $F_{\tilde{\mu}_x}(t) = \mu(\mathbb{B}(x,t))$ and $F_{\tilde{\mu}_x}^{-1}(m) = \delta_{\mu,m}(x)$

$$\int_{\mathbb{R}^d} \|h - x\|^2 d\tilde{\mu}(h) \geq \int_0^{m_0} F_{\mu_x}^{-1}(m)^2 dm = \int_0^{m_0} \delta_{\mu,m}(x)^2 dm$$

Another expression for d_{μ,m_0}

$$d_{\mu,m_0}(x) = \min_{\tilde\mu} \left\{ d_W\left(\delta_x, \frac{1}{m_0}\tilde\mu\right) : \tilde\mu(\mathbb{R}^d) = m_0 \text{ and } \tilde\mu \leq \mu \right\}$$

Proof:

$$\int_{\mathbb{R}^d} \|h-x\|^2 d\tilde\mu(h) = \int_{\mathbb{R}_+} t^2 d\tilde\mu_x(t) = \int_0^{m_0} F_{\tilde\mu_x}^{-1}(m)^2 dm$$

pushforward of $\tilde\mu$ by the distance function to x.

$F_{\tilde\mu_x}(t) = \tilde\mu_x([0,t))$ is the cumulative function of $\tilde\mu_x$ and $F_{\tilde\mu_x}^{-1}(m) = \inf\{t \in \mathbb{R} : F_{\tilde\mu_x}(t) > m\}$ is its generalized inverse

Equality iff $F_{\tilde\mu_x}^{-1}(m) = F_{\mu_x}^{-1}(m)$ for almost every m
$\Rightarrow$ equality if $\tilde\mu = \tilde\mu_{x,m_0}$

$$\int_{\mathbb{R}^d} \|h-x\|^2 d\tilde\mu(h) \geq \int_0^{m_0} F_{\mu_x}^{-1}(m)^2 dm = \int_0^{m_0} \delta_{\mu,m}(x)^2 dm$$

Semiconcavity of d^2_{μ,m_0}

Theorem: Let μ be a probability measure in $\mathbb{R}^d$ and let $m_0 \in (0,1)$.

1. d^2_{μ,m_0} is 1-semiconcave, i.e. x $\in \mathbb{R}^d \mapsto \|x\|^2 - d^2_{\mu,m_0}$ is convex.
2. d^2_{μ,m_0} is differentiable almost everywhere in $\mathbb{R}^d$, with gradient defined by
$$\nabla_x d^2_{\mu,m_0} = \frac{2}{m_0} \int_{h \in \mathbb{R}^d} (x-h)\, d\tilde{\mu}_{x,m_0}(h)$$
3. the function $x \in \mathbb{R}^d \mapsto d_{\mu,m_0}(x)$ is 1-Lipschitz.

Example. Let $C = \{p_1, \ldots, p_n\}$ and $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{p_i}$. Let $p_C^k(x)$ denote the kth nearest neighbor to x in C, and set $m_0 = k_0/n$:

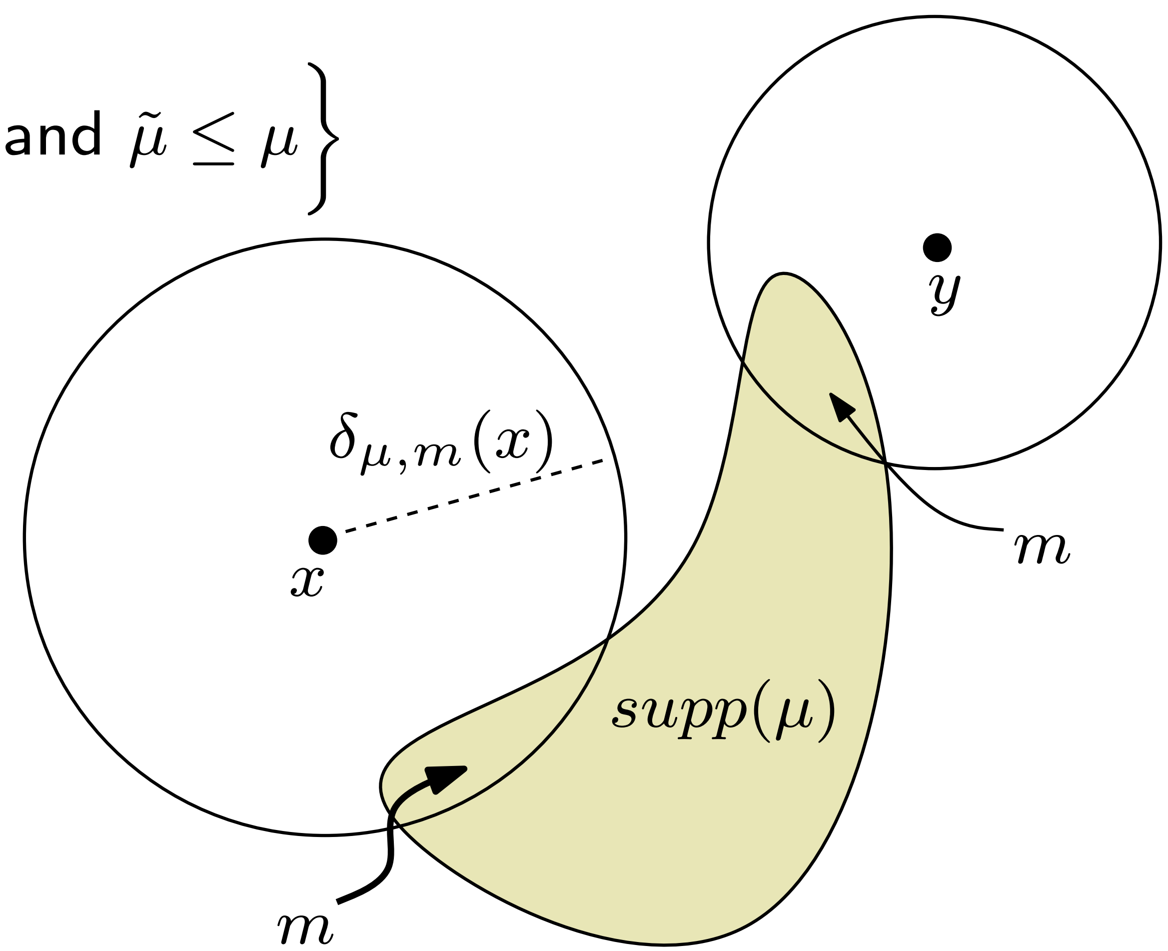
$$\nabla d^2_{\mu,m_0}(x) = 2 d_{\mu,m_0} \nabla d_{\mu,m_0} = \frac{2}{k_0} \sum_{k=1}^{k_0} (x - p_C^k(x))$$

Semiconcavity of d^2_{μ,m_0}

Proof:

$$d^2_{\mu,m_0}(y) = \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \|y-h\|^2\, d\tilde{\mu}_{y,m_0}(h)$$
$$\leq \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \|y-h\|^2\, d\tilde{\mu}_{x,m_0}(h)$$

$$d_{\mu,m_0}(x) = \min_{\tilde{\mu}} \left\{ d_W\left(\delta_x, \frac{1}{m_0}\tilde{\mu}\right) : \tilde{\mu}(\mathbb{R}^d) = m_0 \text{ and } \tilde{\mu} \leq \mu \right\}$$



Semiconcavity of d^2_{μ,m_0}

Proof:

$$\begin{aligned}
d^2_{\mu,m_0}(y) &= \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \|y-h\|^2\, d\tilde{\mu}_{y,m_0}(h) \\
&\leq \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \|y-h\|^2\, d\tilde{\mu}_{x,m_0}(h) \\
&= \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \left(\|x-h\|^2 + 2\langle x-h, y-x\rangle + \|y-x\|^2\ \right)\, d\tilde{\mu}_{x,m_0}(h) \\
&= d^2_{\mu,m_0}(x) + \|y-x\|^2 + \langle V, y-x\rangle
\end{aligned}$$

with $V = \frac{2}{m_0}\int_{h\in\mathbb{R}^d}[x-h]\, d\tilde{\mu}_{x,m_0}(h)$.

Semiconcavity of d^2_{μ,m_0}

Proof:

$$\begin{aligned}
d^2_{\mu,m_0}(y) &= \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \|y-h\|^2\, d\tilde{\mu}_{y,m_0}(h) \\
&\leq \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \|y-h\|^2\, d\tilde{\mu}_{x,m_0}(h) \\
&= \frac{1}{m_0}\int_{h\in\mathbb{R}^d} \left(\|x-h\|^2 + 2\langle x-h, y-x\rangle + \|y-x\|^2\right)\, d\tilde{\mu}_{x,m_0}(h) \\
&= d^2_{\mu,m_0}(x) + \|y-x\|^2 + \langle V, y-x\rangle
\end{aligned}$$

with $V = \frac{2}{m_0}\int_{h\in\mathbb{R}^d}[x-h]\, d\tilde{\mu}_{x,m_0}(h)$.

$$\Rightarrow (\|y\|^2 - d^2_{\mu,m_0}(y)) - (\|x\|^2 - d^2_{\mu,m_0}(x)) \geq \langle 2x - V, x-y\rangle$$

This is the gradient!

Stability of of $\mu \rightarrow d_{\mu,m_0}$

Theorem: If μ and ν are two probability measures on $\mathbb{R}^d$ and $m_0 > 0$, then $\|d_{\mu,m_0} - d_{\nu,m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} d_W(\mu, \nu)$.

Stability of of $\mu \longrightarrow d_{\mu,m_0}$

Theorem: If μ and ν are two probability measures on $\mathbb{R}^d$ and $m_0 > 0$, then $\|d_{\mu,m_0} - d_{\nu,m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} d_W(\mu,\nu)$.

Proof: for $x \in \mathbb{R}^d$, let $\mu^x = (d_x)_\sharp \mu$, $\nu^x = (d_x)_\sharp \nu$ where $d_x : y \in \mathbb{R}^d \mapsto \|y - x\|$.

- $d_W(\mu^x, \nu^x) = \|F_{\mu^x}^{-1} - F_{\nu^x}^{-1}\|_{L^2([0,1])}$

Classical result for measures on $\mathbb{R}$

Stability of of $\mu \to d_{\mu,m_0}$

Theorem: If μ and ν are two probability measures on $\mathbb{R}^d$ and $m_0 > 0$, then $\|d_{\mu,m_0} - d_{\nu,m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} d_W(\mu,\nu)$.

Proof: for $x \in \mathbb{R}^d$, let $\mu^x = (d_x)_\sharp \mu$, $\nu^x = (d_x)_\sharp \nu$ where $d_x : y \in \mathbb{R}^d \mapsto \|y - x\|$.

- $d_W(\mu^x, \nu^x) = \|F_{\mu^x}^{-1} - F_{\nu^x}^{-1}\|_{L^2([0,1])}$ $+ \ \delta_{\mu^x,m}(0) = F_{\mu^x}^{-1}(m)$
- $\left| \sqrt{\int_0^{m_0} \delta^2_{\mu^x,m}(0) dm} - \sqrt{\int_0^{m_0} \delta^2_{\nu^x,m}(0) dm} \right| \leq d_W(\mu^x, \nu^x)$

Stability of of $\mu \to d_{\mu,m_0}$

Theorem: If μ and ν are two probability measures on $\mathbb{R}^d$ and $m_0 > 0$, then $\|d_{\mu,m_0} - d_{\nu,m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} d_W(\mu,\nu)$.

Proof: for $x \in \mathbb{R}^d$, let $\mu^x = (d_x)_\sharp \mu$, $\nu^x = (d_x)_\sharp \nu$ where $d_x : y \in \mathbb{R}^d \mapsto \|y - x\|$.

- $d_W(\mu^x, \nu^x) = \|F_{\mu^x}^{-1} - F_{\nu^x}^{-1}\|_{L^2([0,1])}$
- $\left| \sqrt{\int_0^{m_0} \delta^2_{\mu^x,m}(0)dm} - \sqrt{\int_0^{m_0} \delta^2_{\nu^x,m}(0)dm} \right| \leq d_W(\mu^x, \nu^x)$
- $d_W(\mu^x, \nu^x) \leq d_W(\mu, \nu)$ ⟵ Any transport plan π between μ and ν induces a transport plan $\pi_x = (d_x, d_x)_\sharp \pi$ between μ^x and ν^x

Stability of of $\mu \to d_{\mu,m_0}$

Theorem: If μ and ν are two probability measures on $\mathbb{R}^d$ and $m_0 > 0$, then $\|d_{\mu,m_0} - d_{\nu,m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} d_W(\mu,\nu)$.

Proof: for $x \in \mathbb{R}^d$, let $\mu^x = (d_x)_\sharp \mu$, $\nu^x = (d_x)_\sharp \nu$ where $d_x : y \in \mathbb{R}^d \mapsto \|y - x\|$.

- $d_W(\mu^x, \nu^x) = \|F^{-1}_{\mu^x} - F^{-1}_{\nu^x}\|_{L^2([0,1])}$
- $\left| \sqrt{\int_0^{m_0} \delta^2_{\mu^x,m}(0) dm} - \sqrt{\int_0^{m_0} \delta^2_{\nu^x,m}(0) dm} \right| \leq d_W(\mu^x, \nu^x)$
- $d_W(\mu^x, \nu^x) \leq d_W(\mu, \nu)$
- $|d_{\mu,m_0}(x) - d_{\nu,m_0}(x)| \leq \frac{1}{\sqrt{m_0}} d_W(\mu^x, \nu^x) \leq \frac{1}{\sqrt{m_0}} d_W(\mu, \nu)$

To summarize

Theorem [C-Cohen-Steiner-Mérigot'09]

1. the function $x \mapsto d_{\mu,m_0}(x)$ is 1-Lipschitz;
2. the function $x \mapsto \|x\|^2 - d^2_{\mu,m_0}(x)$ is convex;
3. the map $\mu \mapsto d_{\mu,m_0}$ from probability measures to continuous functions is $\frac{1}{\sqrt{m_0}}$–Lipschitz, ie

$$\|d_{\mu,m_0} - d_{\mu',m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} d_W(\mu, \mu')$$

In practice: d_{μ,m_0} and $\nabla d_{\mu,m_0}$ are very easy to compute for $\mu = \sum_{i=1}^n \delta_{p_i}$, $C = \{p_1, \cdots p_n\} \subset \mathbb{R}^d$!

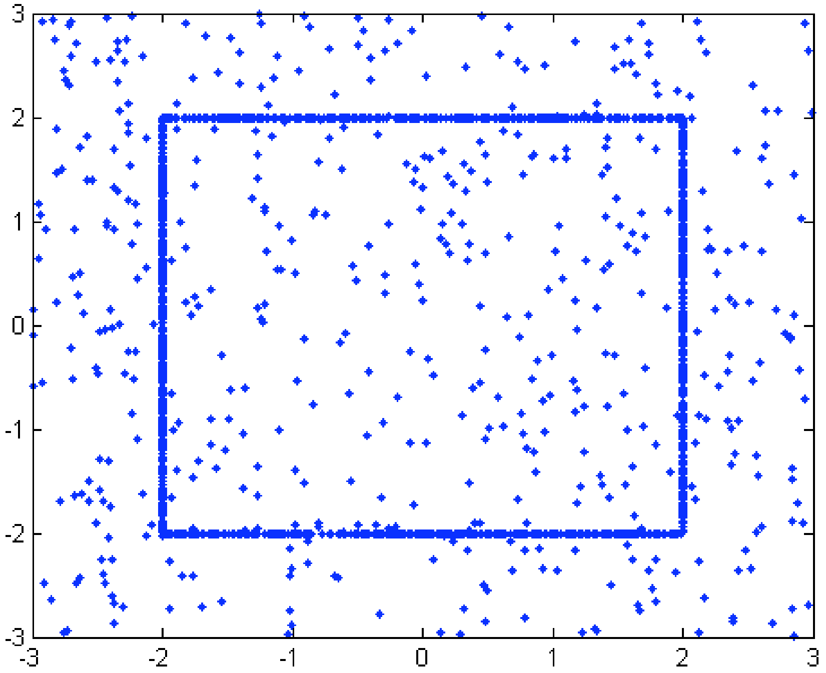
Consequences

Most of the topological and geometric inference for distance functions transpose to distance to a measure functions!

$\Longrightarrow$ This gives a way to associate robust geometric features to any probability measure in an Euclidean space:

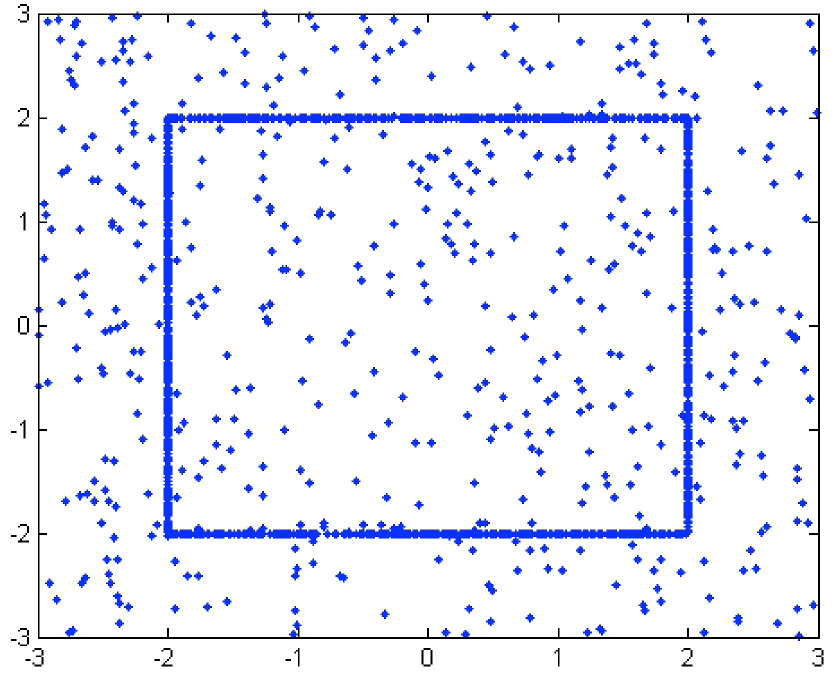
- offsets topology and geometry,
- analogous of the notions of medial axes,
- L^1 stability of $\nabla d_{\mu,m_0}$
- $\cdots$

Example: a square with outliers

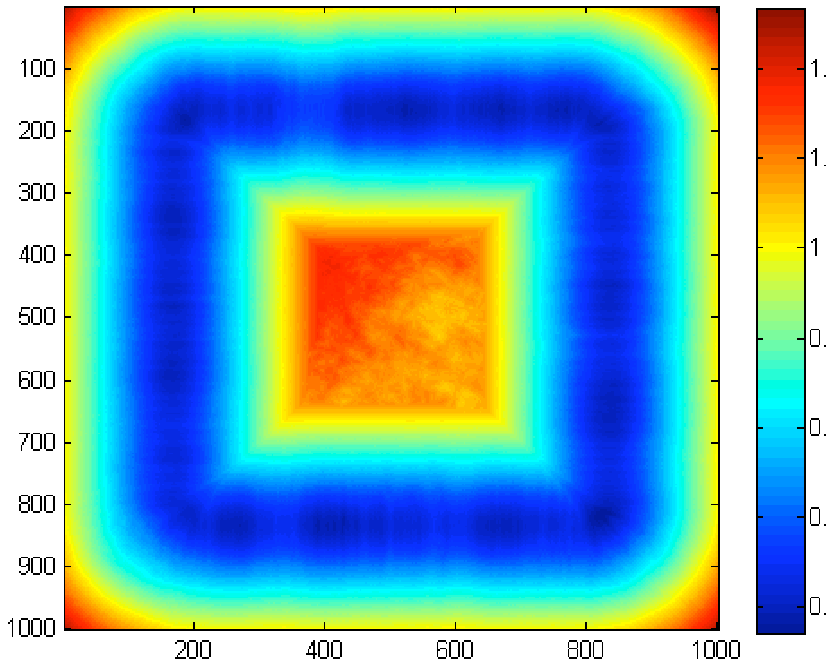


2300 points, 20% outliers

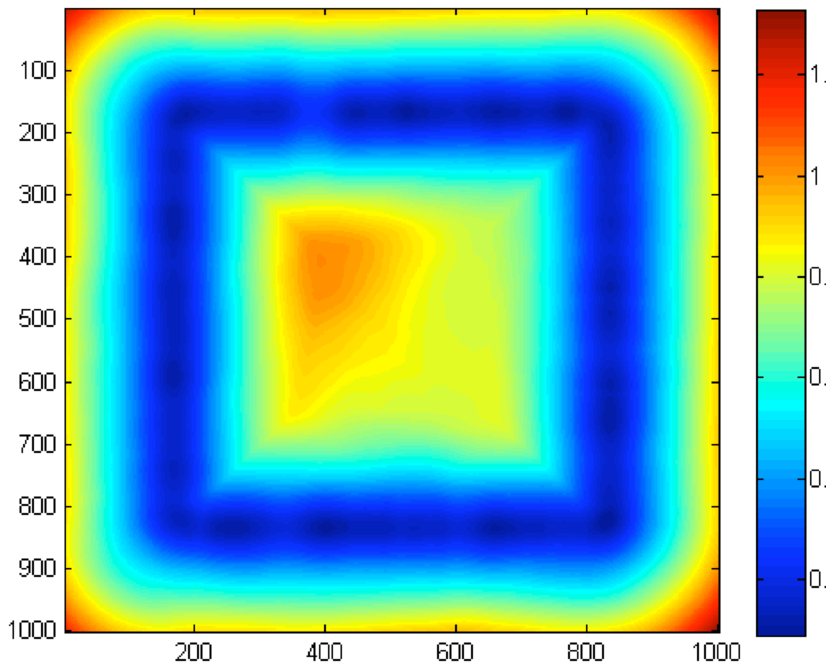
Example: a square with outliers



2300 points, 20% outliers

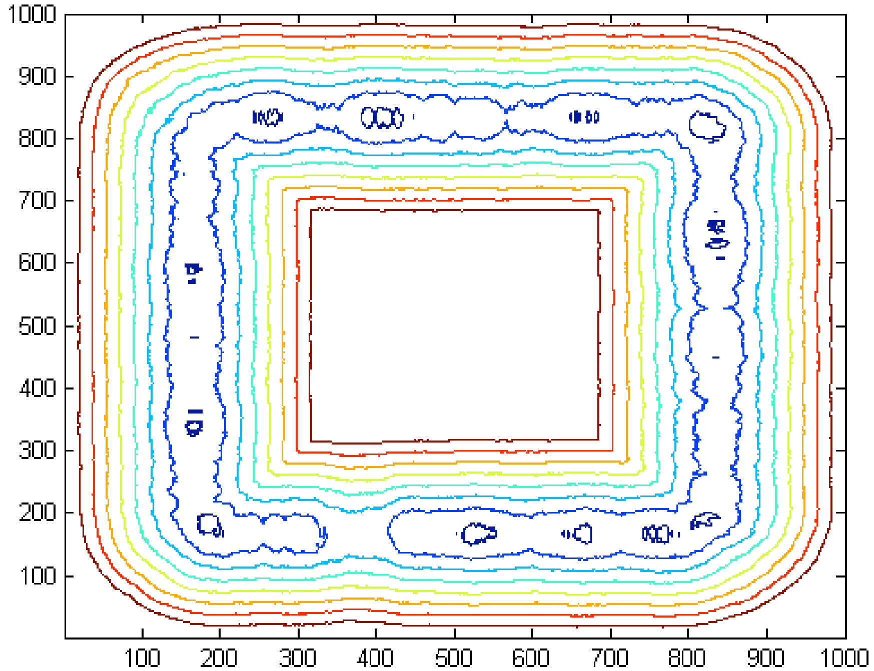


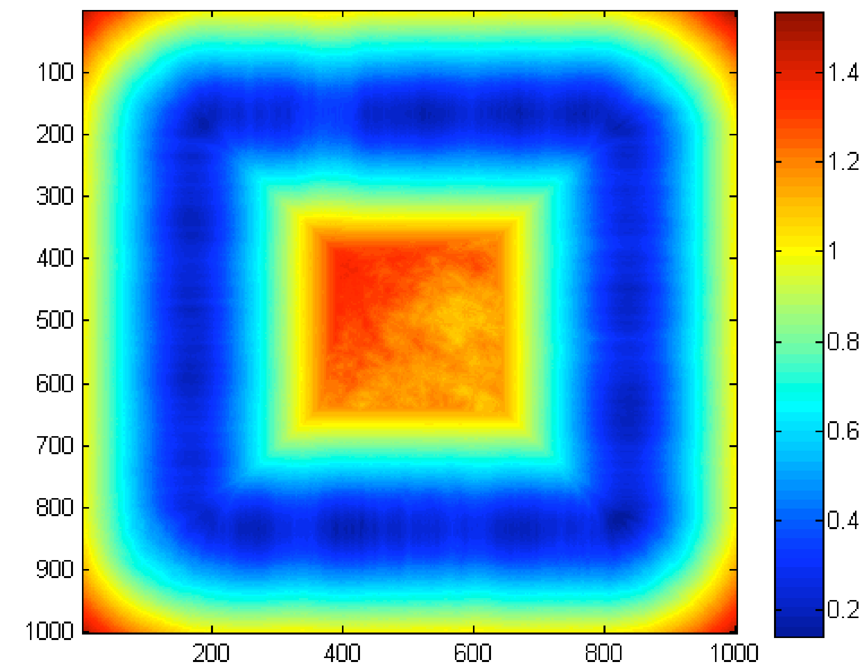
δ_{μ,m_0}, $m_0 = 0.023$ $(k = 50)$



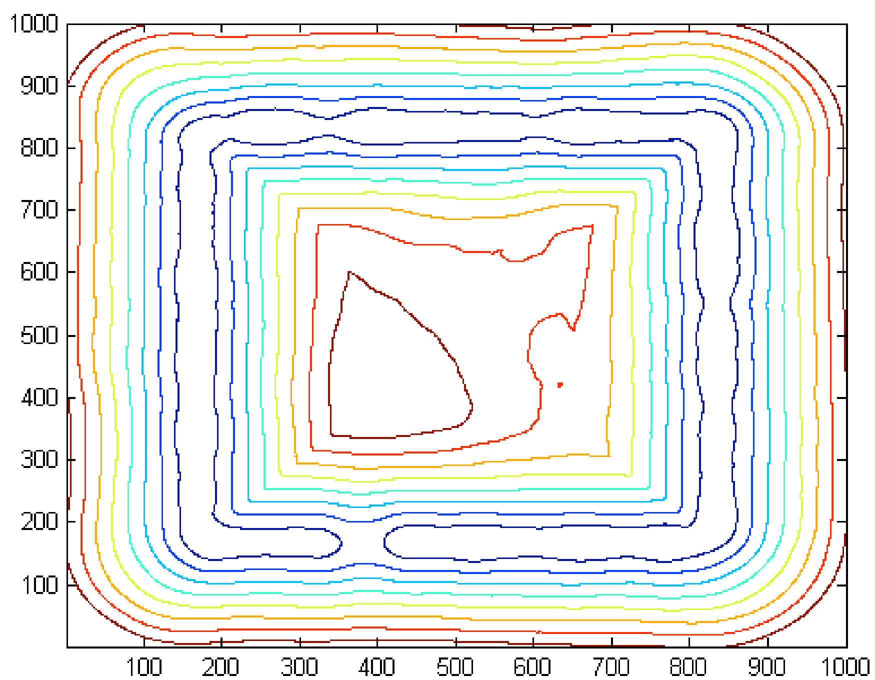
d_{μ,m_0}, $m_0 = 0.023$ $(k = 50)$

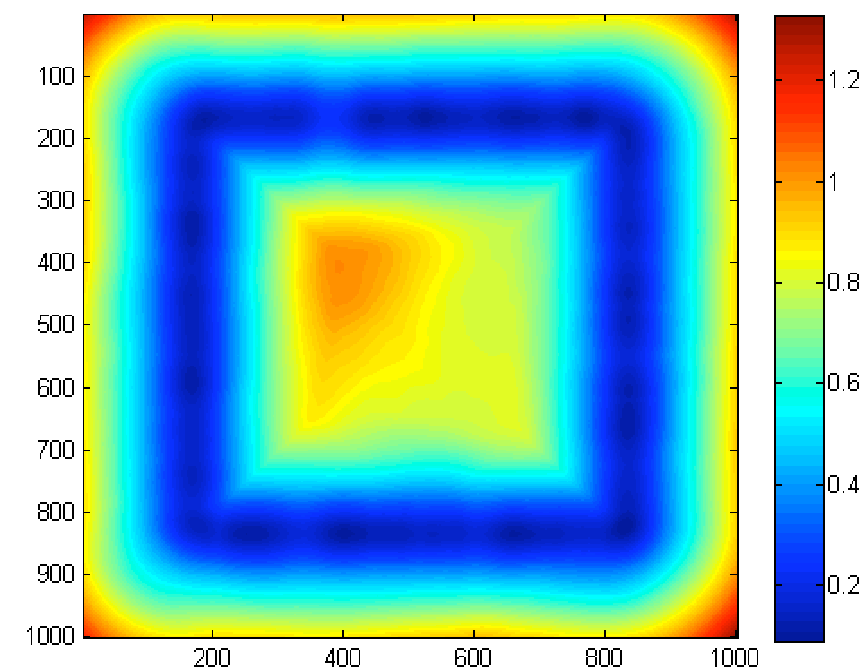
Example: a square with outliers





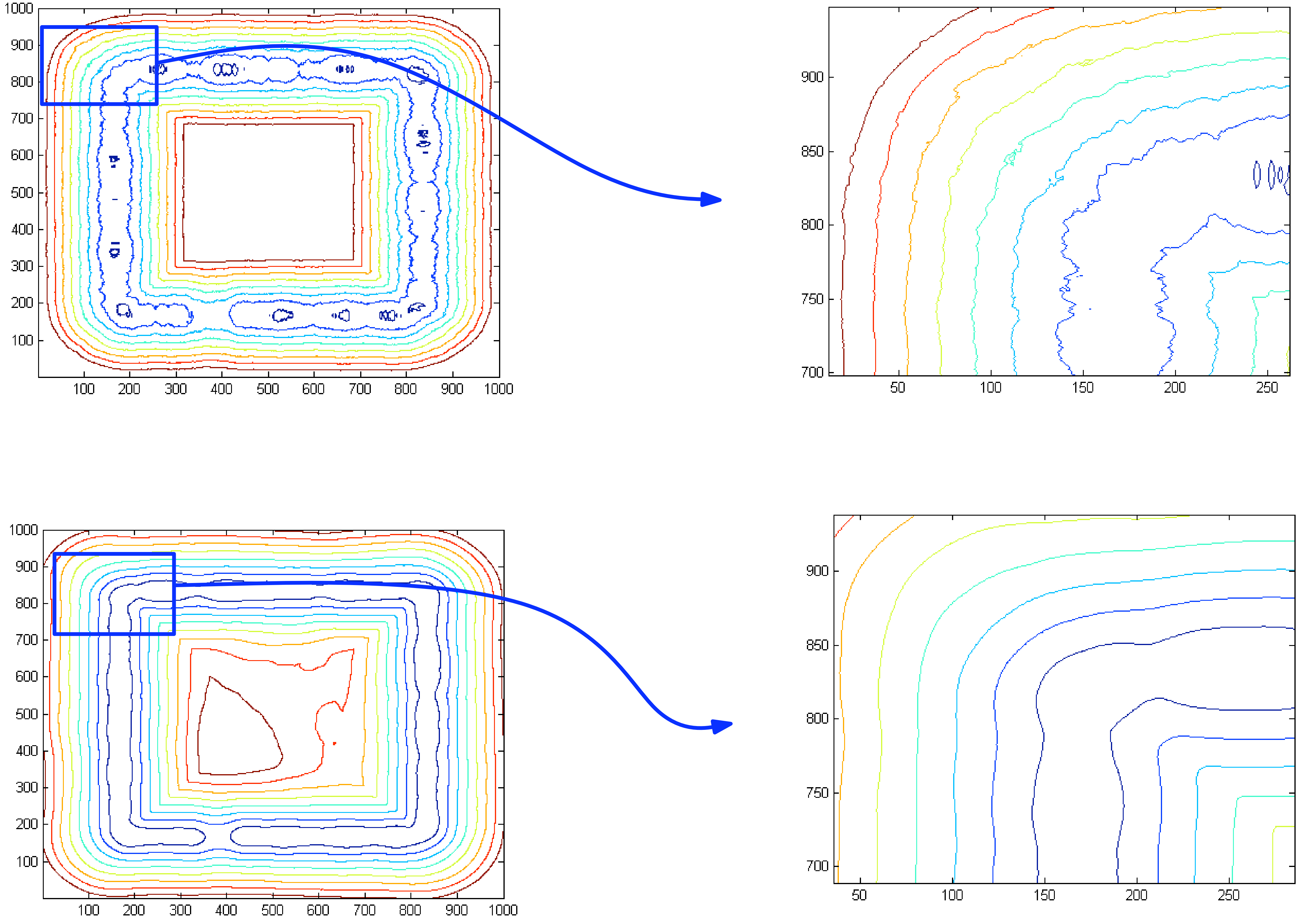
δ_{μ,m_0}, $m_0 = 0.023$ $(k = 50)$



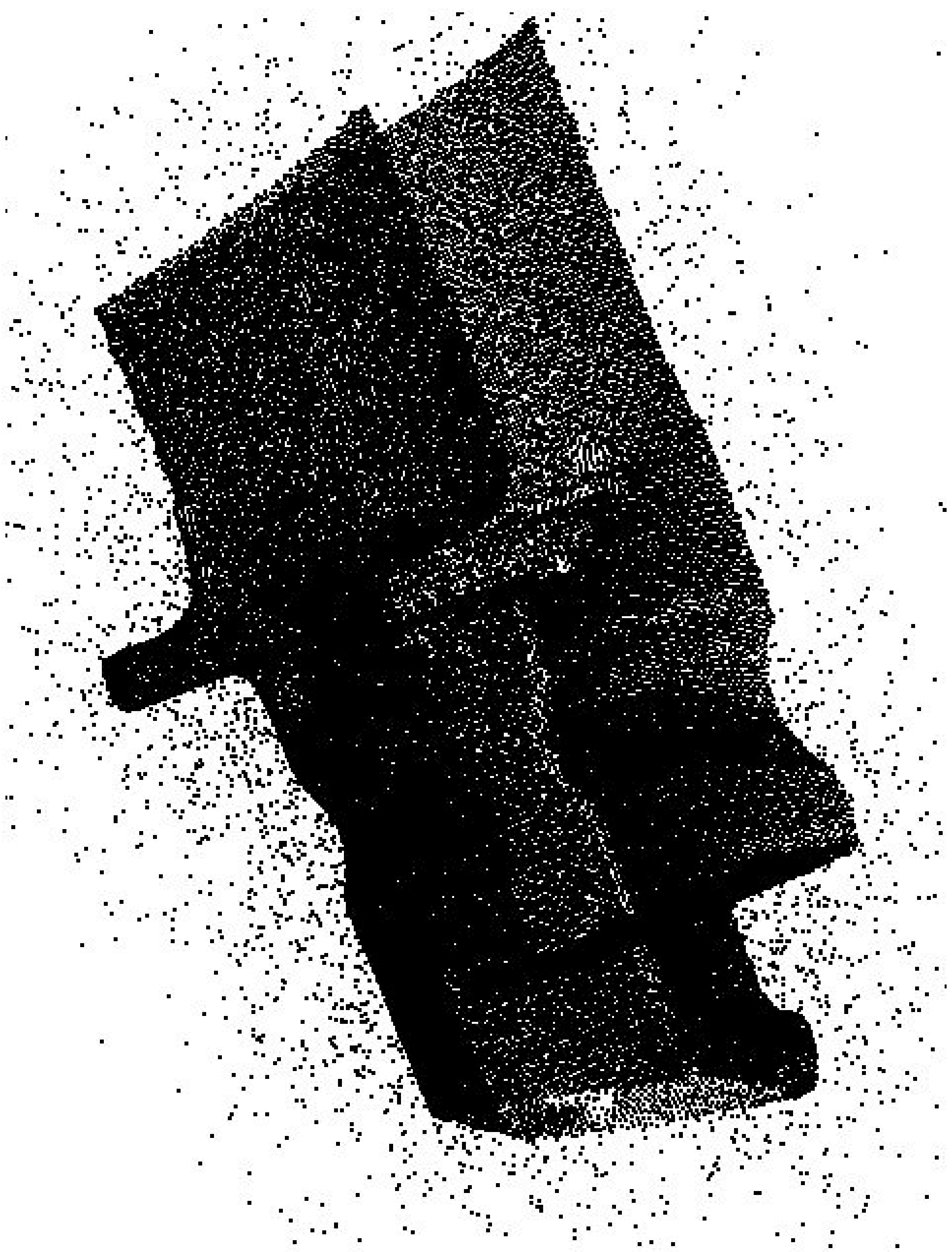


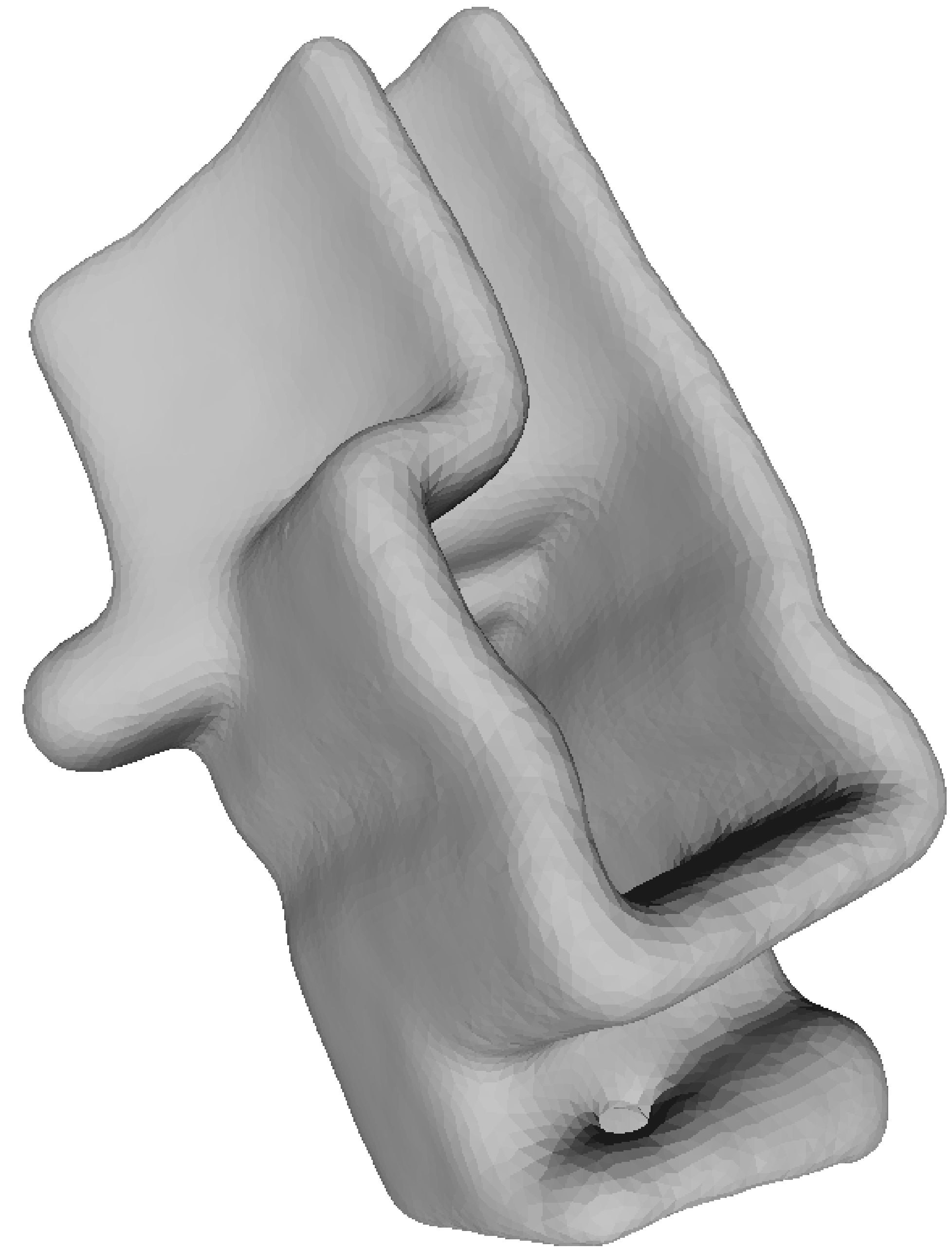
d_{μ,m_0}, $m_0 = 0.023$ $(k = 50)$

Example: a square with outliers



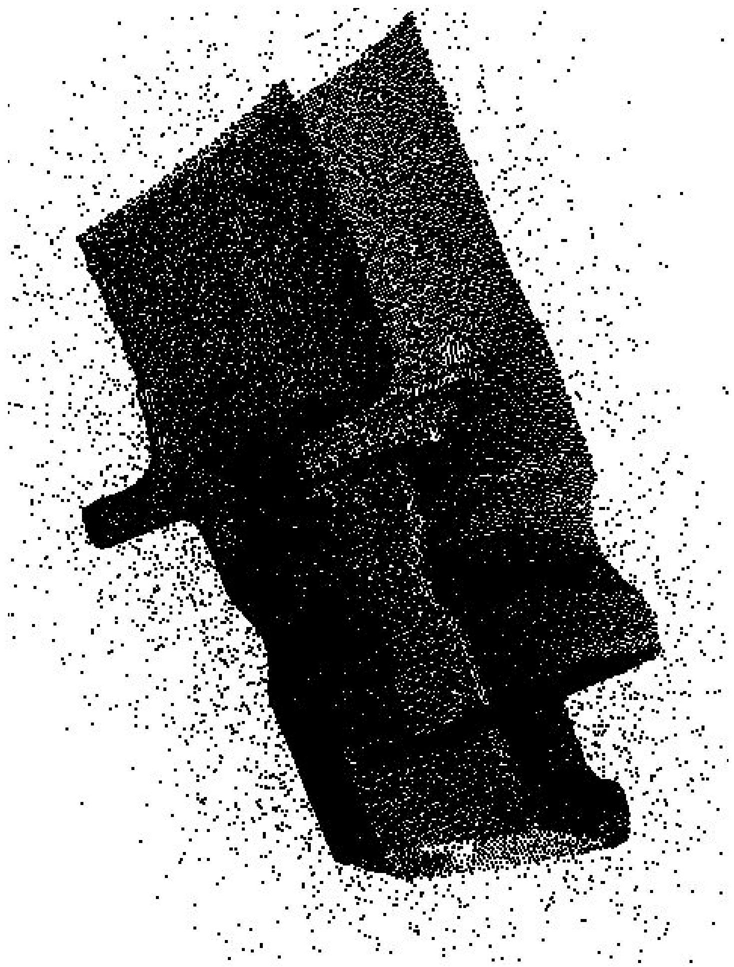
A 3D example

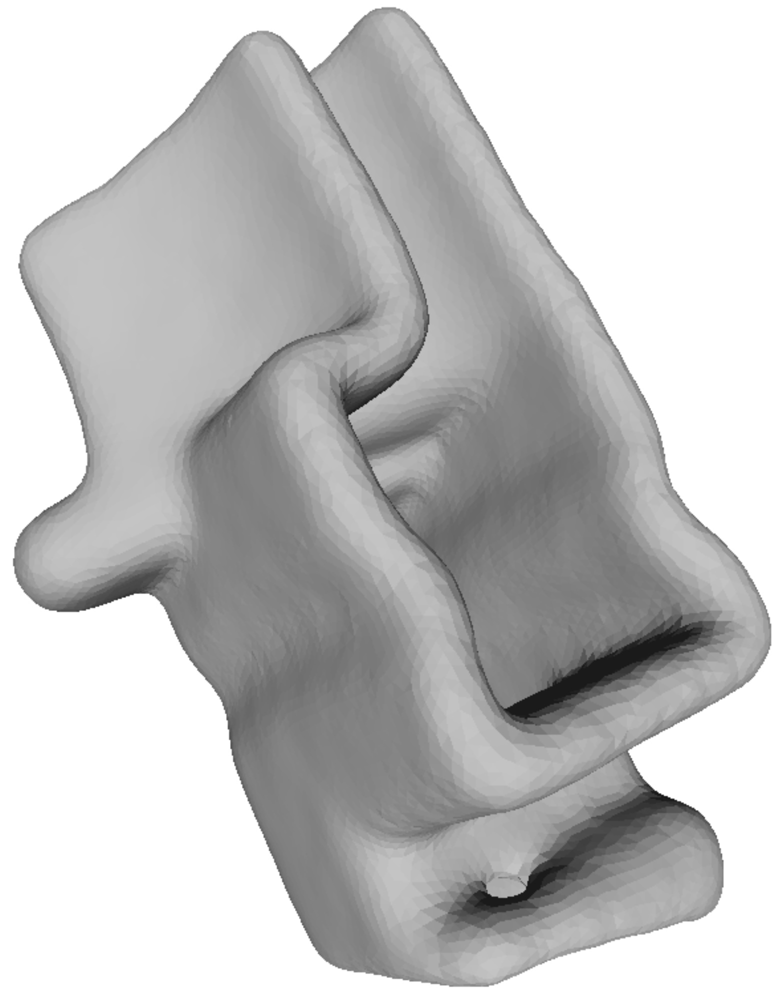




Reconstruction of an offset of a mechanical part from a noisy approximation with 10% outliers

A reconstruction theorem

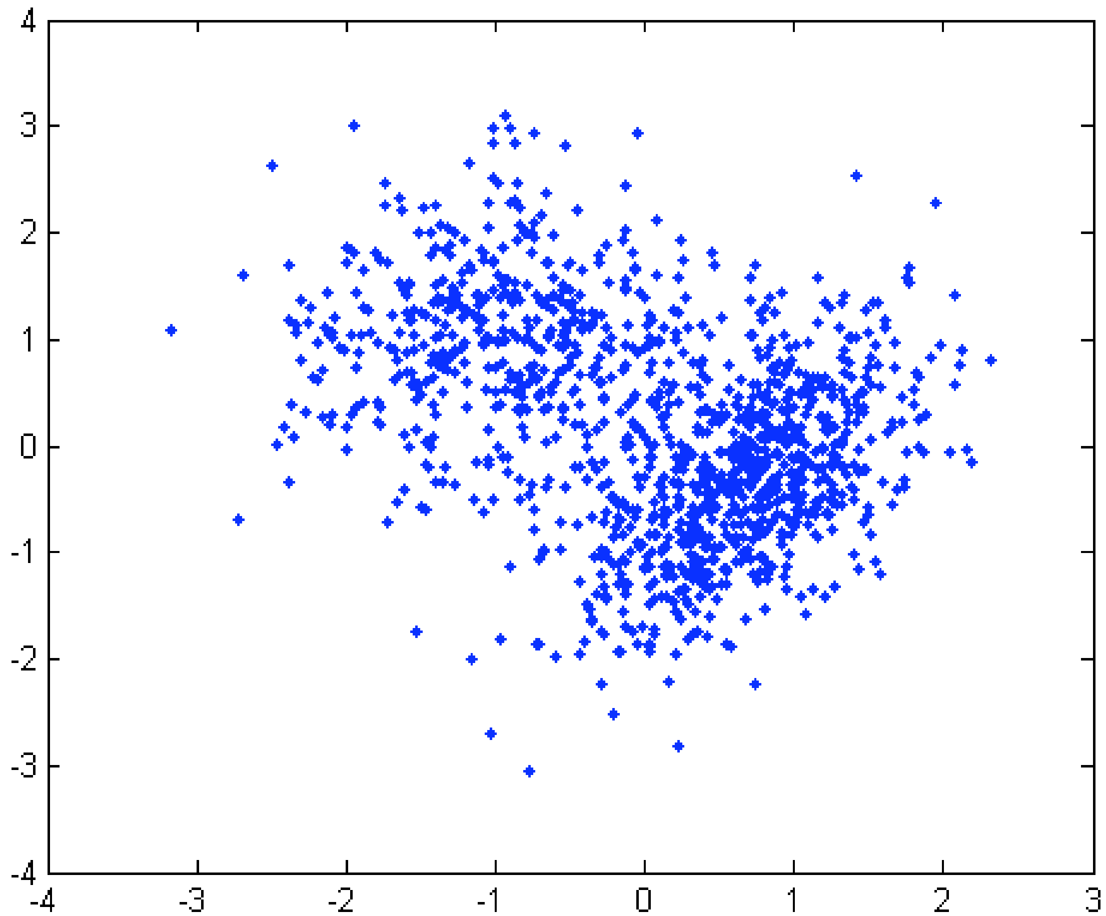




$\exists C > 0$ s.t. $\forall x \in K$, $\mu(\mathbb{B}(x,\varepsilon)) \geq C\varepsilon^k$ as soon as ε is small enough.

Theorem: Let μ be a probability measure *of dimension at most* $k > 0$ with compact support $K \subset \mathbb{R}^d$ such that $r_\alpha(K) > 0$ for some $\alpha \in (0,1]$. For any $0 < \eta < r_\alpha(K)$, there exists positive constants $m_1 = m_1(\mu,\alpha,\eta) > 0$ and $C = C(m_1) > 0$ such that:
for any $m_0 < m_1$ and any probability measure μ' such that $W_2(\mu,\mu') < C\sqrt{m_0}$, the sublevel set $d_{\mu',m_0}^{-1}((-\infty,\eta])$ is isotopic to the offsets $d_K^{-1}([0,r])$ of K for $0 < r < r_\alpha(K)$.

Comparison to kNN density estimation



$$\hat{\mu} = \frac{1}{1200} \sum_{i=1}^{1200} \delta_{p_i}$$

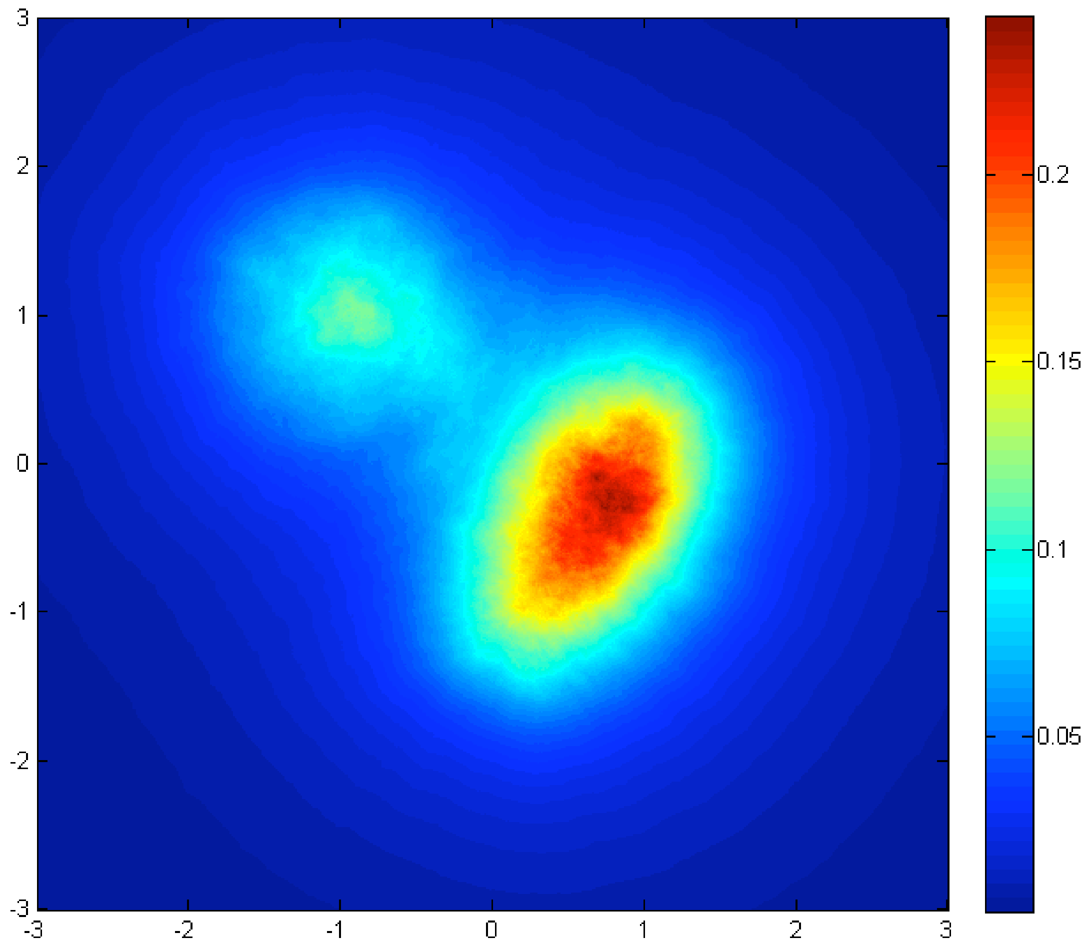
Data: 1200 points $p_1, \cdots, p_{1200}$

Density is estimated using

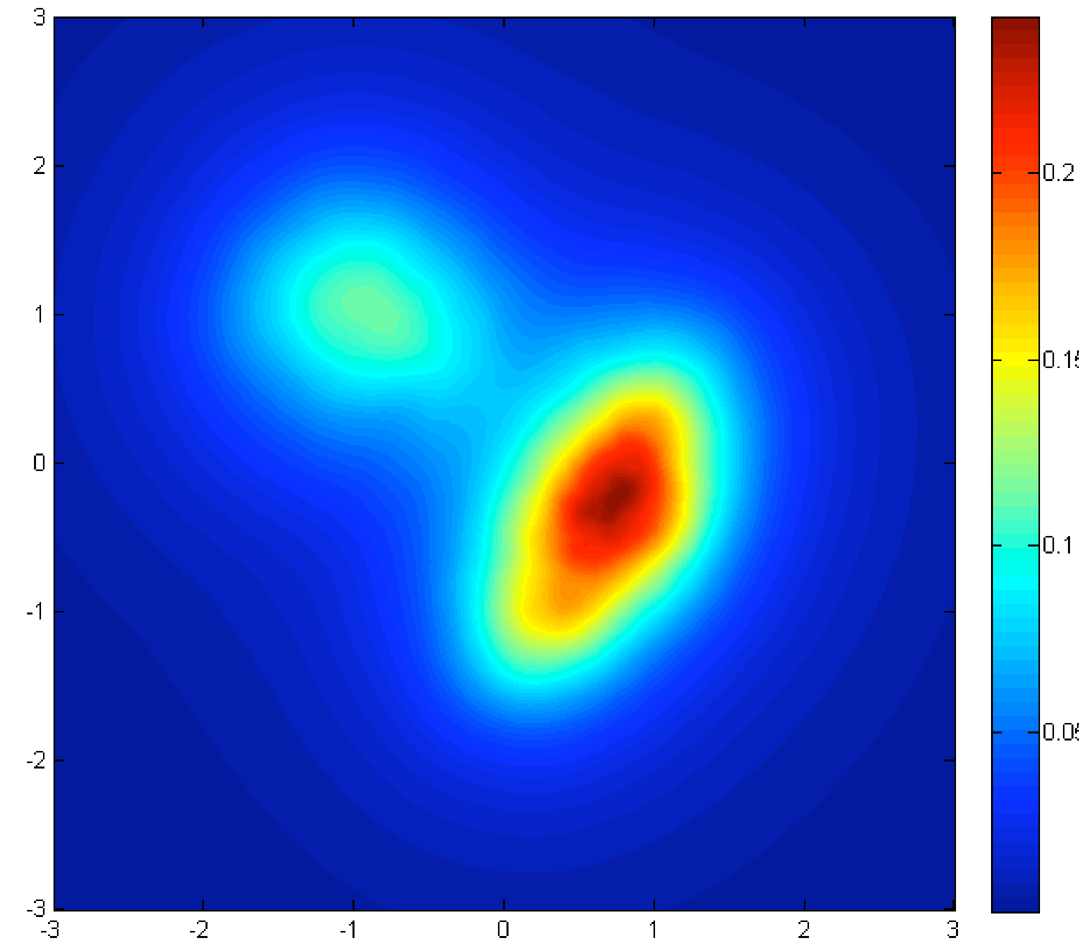
1. $x \mapsto \frac{m_0}{\omega_{d-1}(\delta_{\hat{\mu},m_0}(x))}$, $m_0 = 150/1200$ ($k = 150$) (Devroye-Wagner'77).

2. $\frac{m_0}{2\pi d_{\hat{\mu},m_0}(x)^2}$, $m_0 = 150/1200$ ($k = 150$).

Comparison to kNN density estimation



1.



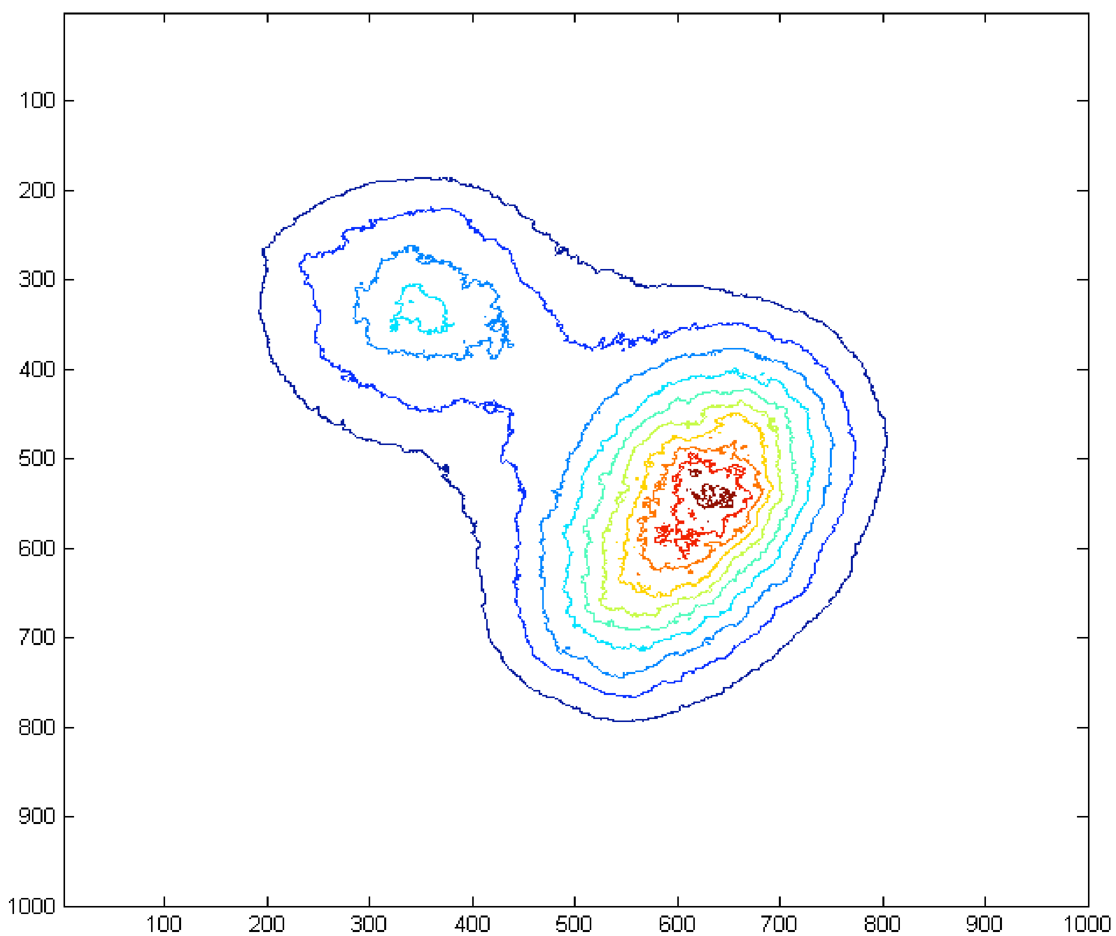
2.

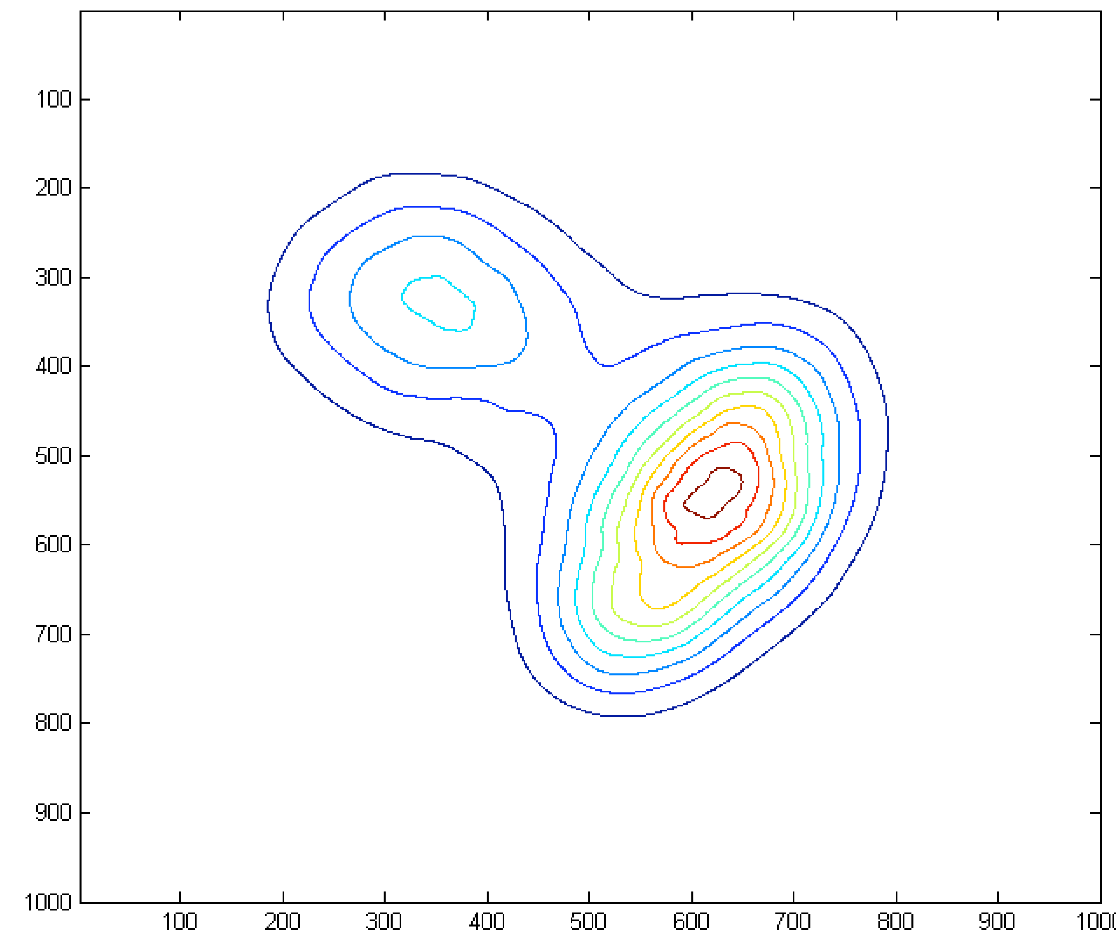
Density is estimated using

1. $x \mapsto \frac{m_0}{\omega_{d-1}(\delta_{\hat{\mu},m_0}(x))}$, $m_0 = 150/1200$ $(k = 150)$ (Devroye-Wagner'77).

2. $\frac{m_0}{2\pi d_{\hat{\mu},m_0}(x)^2}$, $m_0 = 150/1200$ $(k = 150)$.

Comparison to kNN density estimation

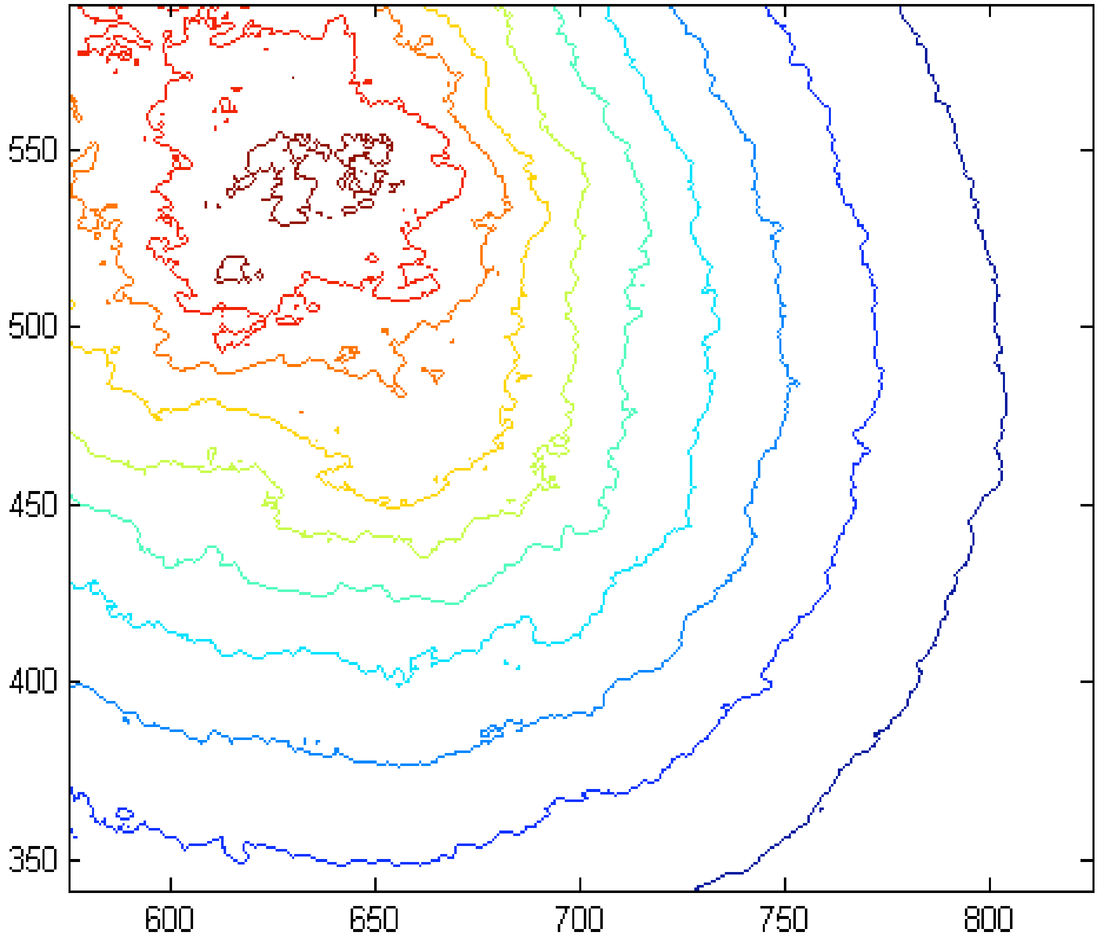




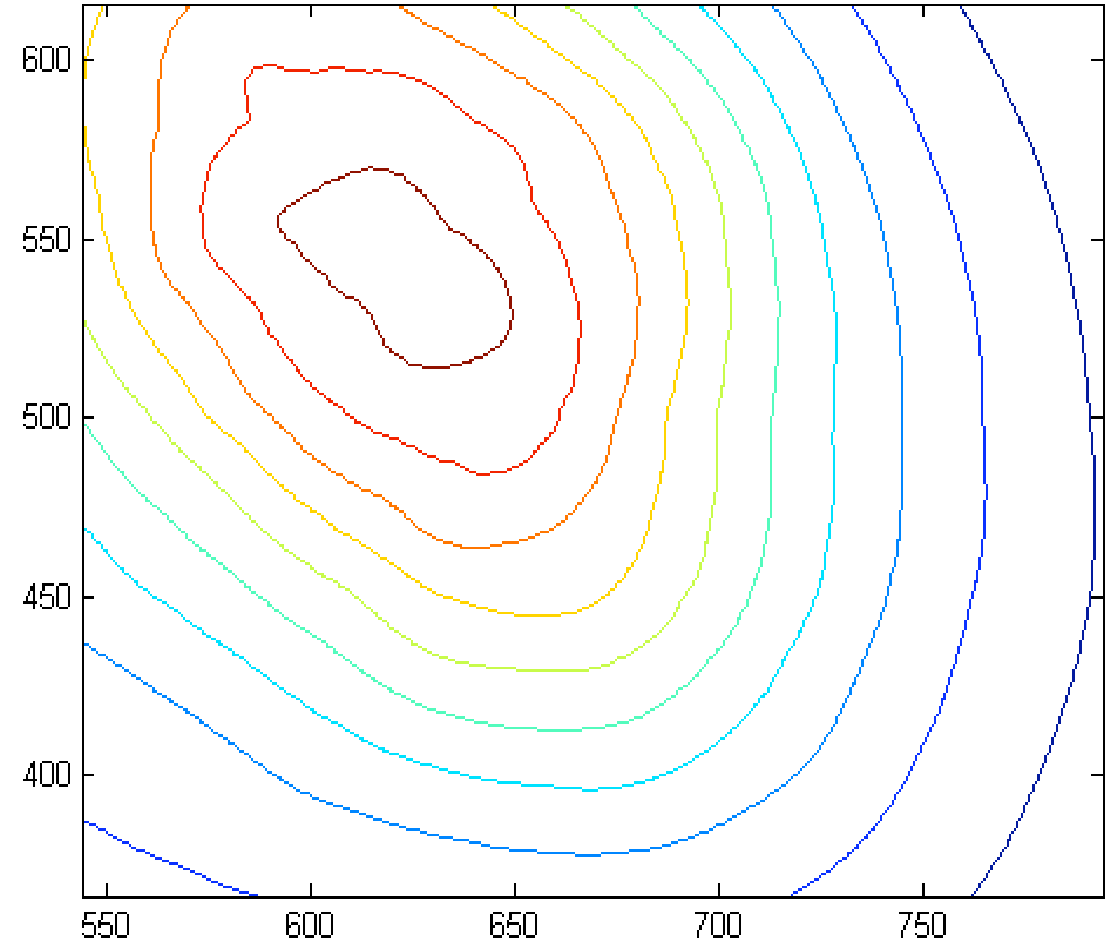
Density is estimated using

1. $x \mapsto \frac{m_0}{\omega_{d-1}(\delta_{\hat{\mu},m_0}(x))}$, $m_0 = 150/1200$ $(k = 150)$ (Devroye-Wagner'77).
2. $\frac{m_0}{2\pi d_{\hat{\mu},m_0}(x)^2}$, $m_0 = 150/1200$ $(k = 150)$.

Comparison to kNN density estimation



1.



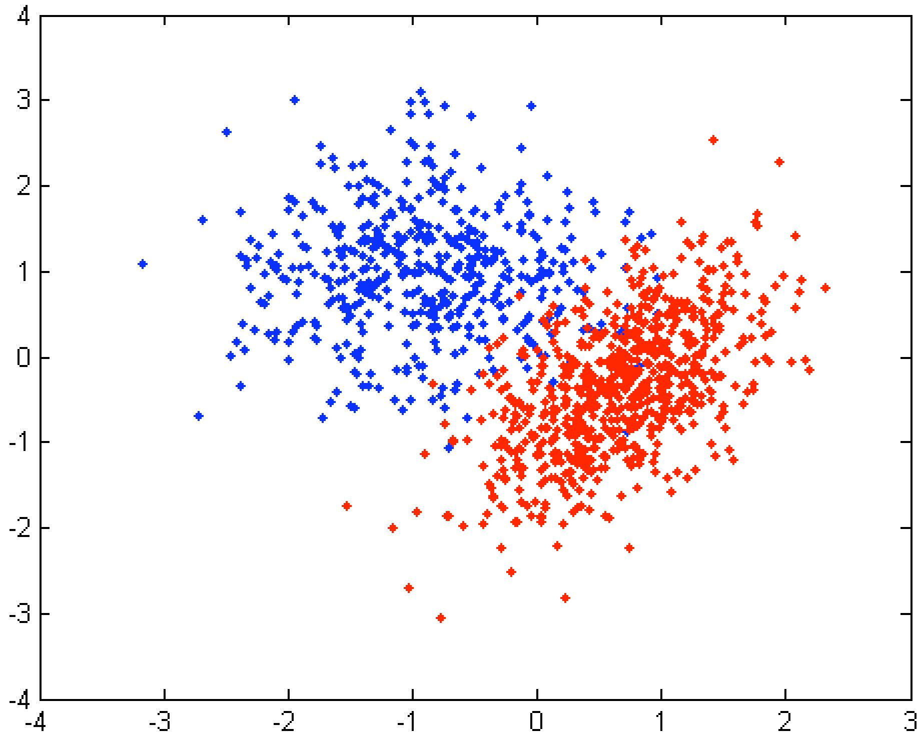
2.

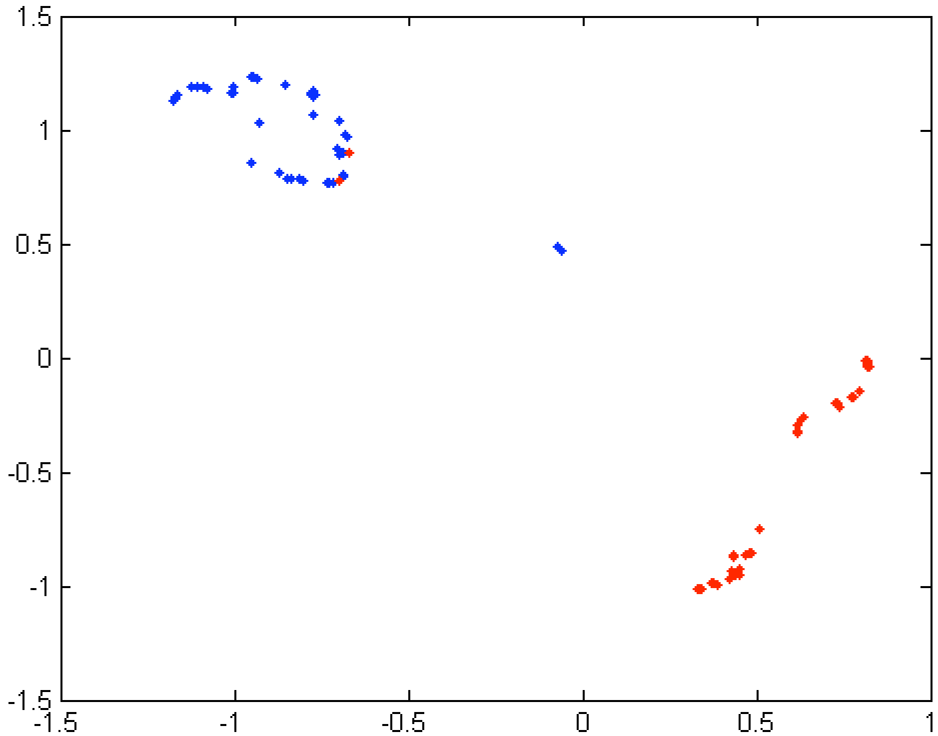
Density is estimated using

1. $x \mapsto \frac{m_0}{\omega_{d-1}(\delta_{\hat{\mu},m_0}(x))}$, $m_0 = 150/1200$ $(k = 150)$ (Devroye-Wagner'77).

2. $\frac{m_0}{2\pi d_{\hat{\mu},m_0}(x)^2}$, $m_0 = 150/1200$ $(k = 150)$.

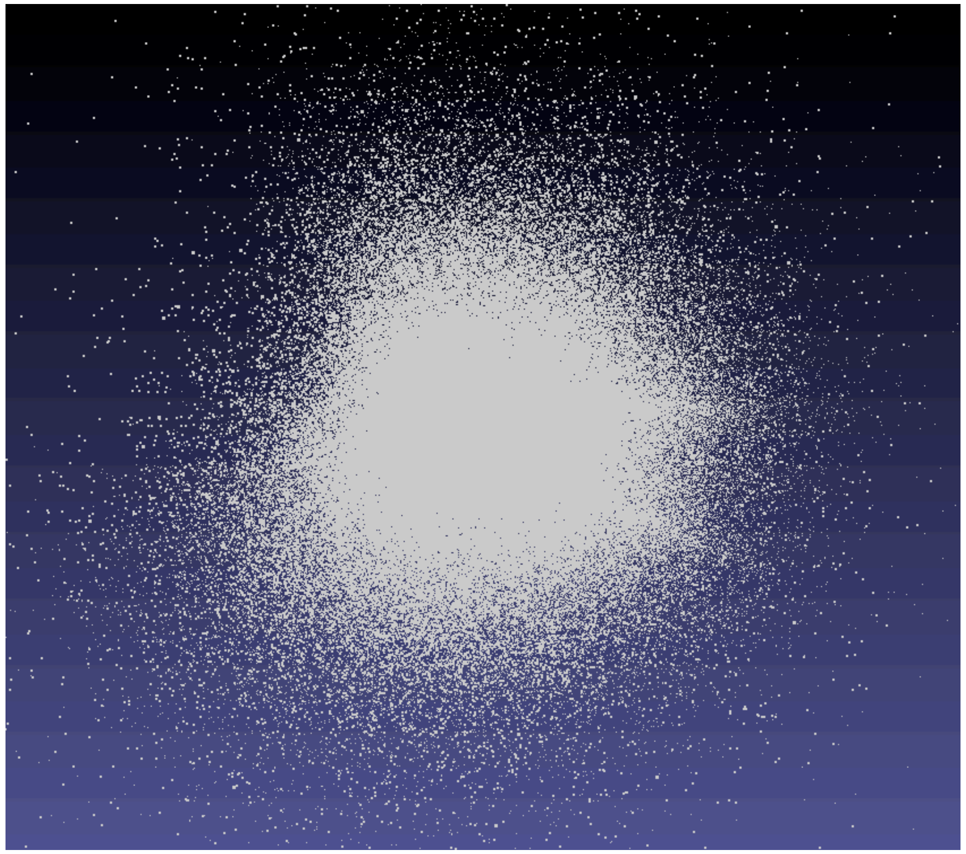
Pushing data along the gradient of d_{μ,m_0}

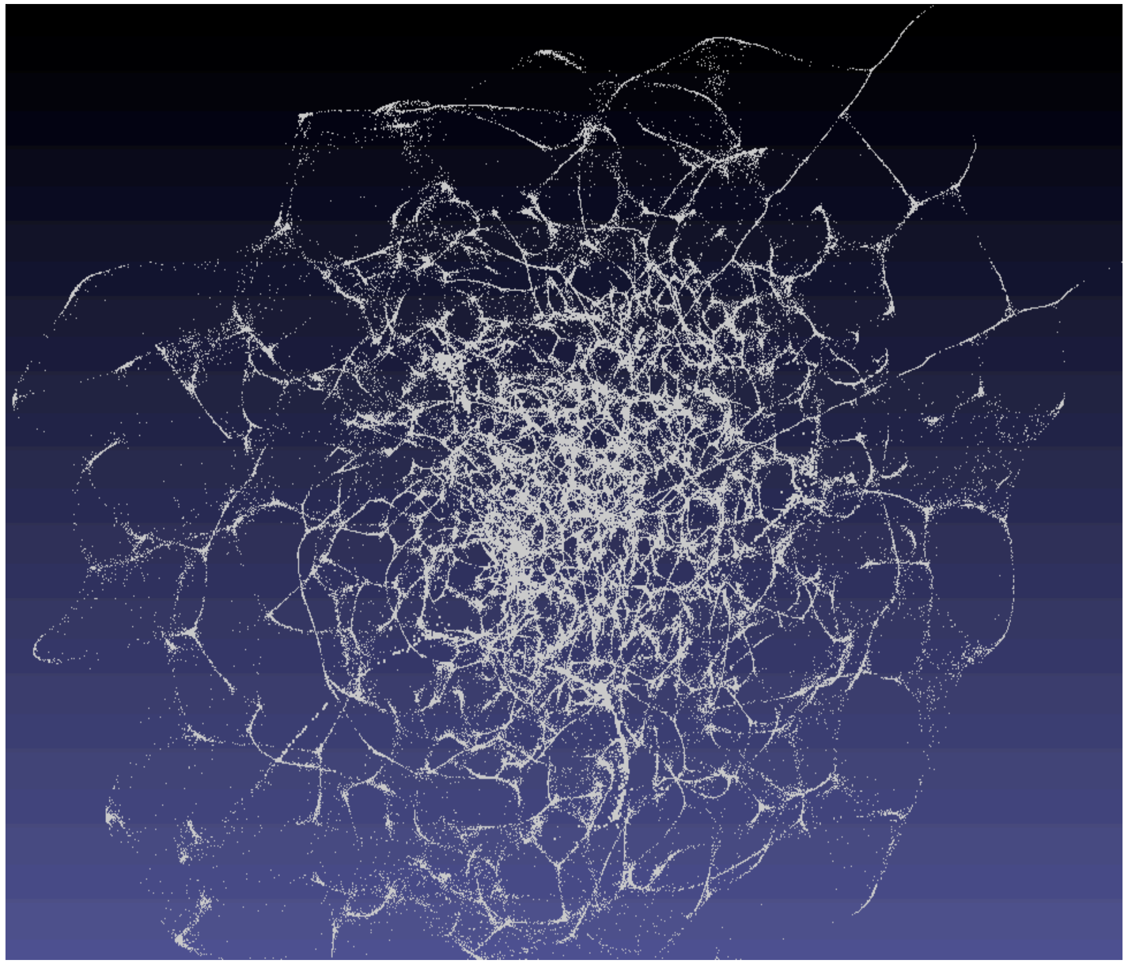


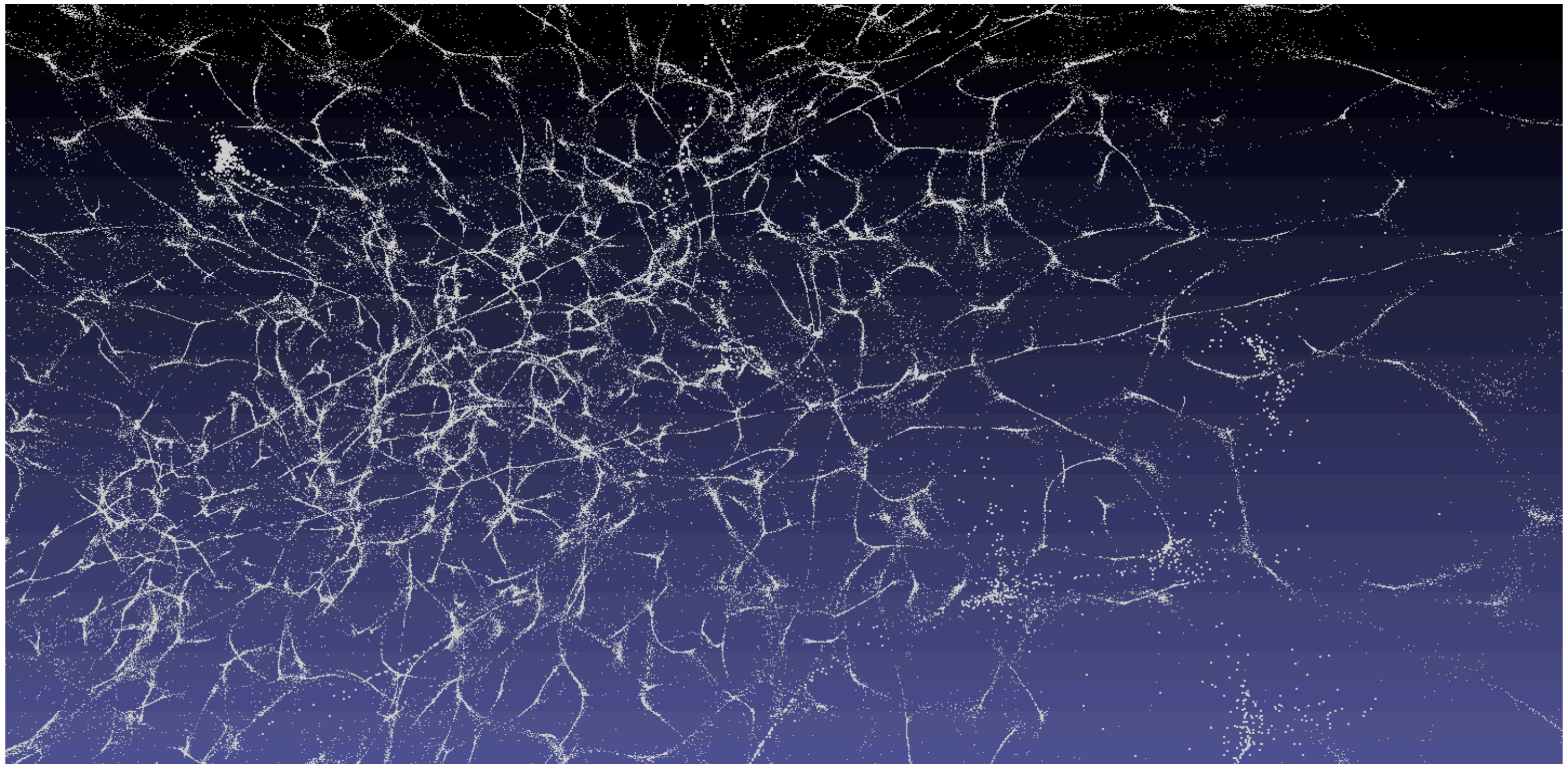


- Mean-Shift like algorithm (Fukunaga-Hostetler'75, Comaniciu-Meer '02)
- Theoretical guarantees on the convergence of the algorithm and "smoothness" of trajectories.

Pushing data along the gradient of d_{μ,m_0}







Take-home messages

- $\mu \mapsto d_{\mu,m_0}$ provide a way to associate geometry to a measure in Euclidean space.
- d_{μ,m_0} is robust to Wasserstein perturbations : outliers and noise are easily handled (no assumption on the nature of the noise).
- d_{μ,m_0} shares regularity properties with the usual distance function to a compact.
- Geometric stability results in this measure-theoretic setting : topology/geometry of the sublevel sets of d_{μ,m_0}, stable notion of persistence diagram for μ,...
- Algorithm: for finite point clouds d_{μ,m_0} and $\nabla(d_{\mu,m_0})$ can be easily and efficiently computed in any dimension.

To get more details:
http://geometrica.saclay.inria.fr/team/Fred.Chazal/papers/RR-6930.pdf