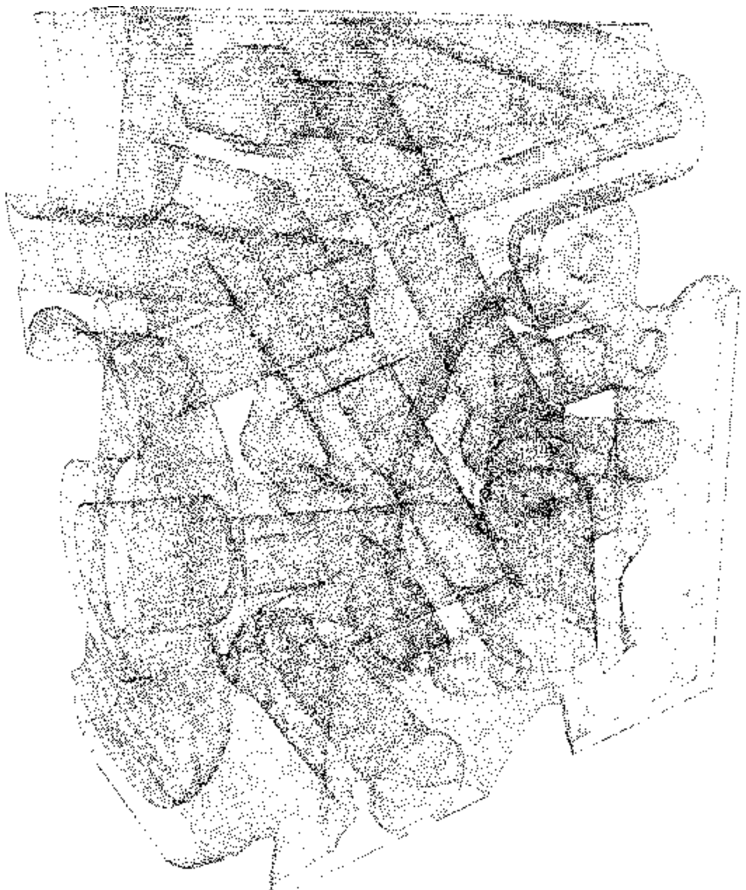
Collège de France
31 mai 2017

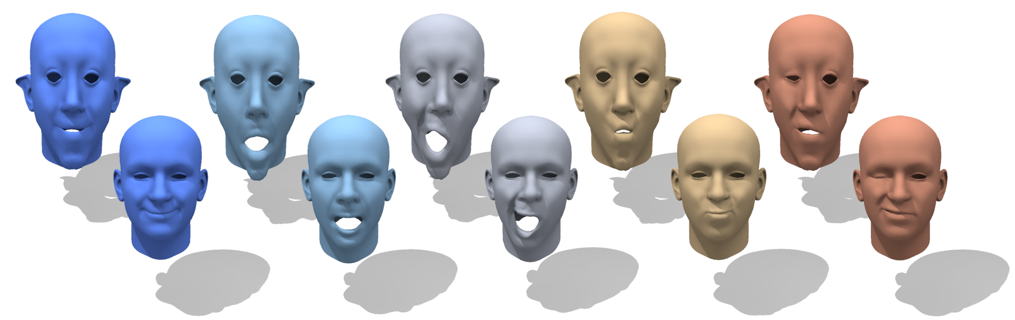
Analyse Topologique des Données

Frédéric Chazal
DataShape team
INRIA Saclay - Ile-de-France
frederic.chazal@inria.fr

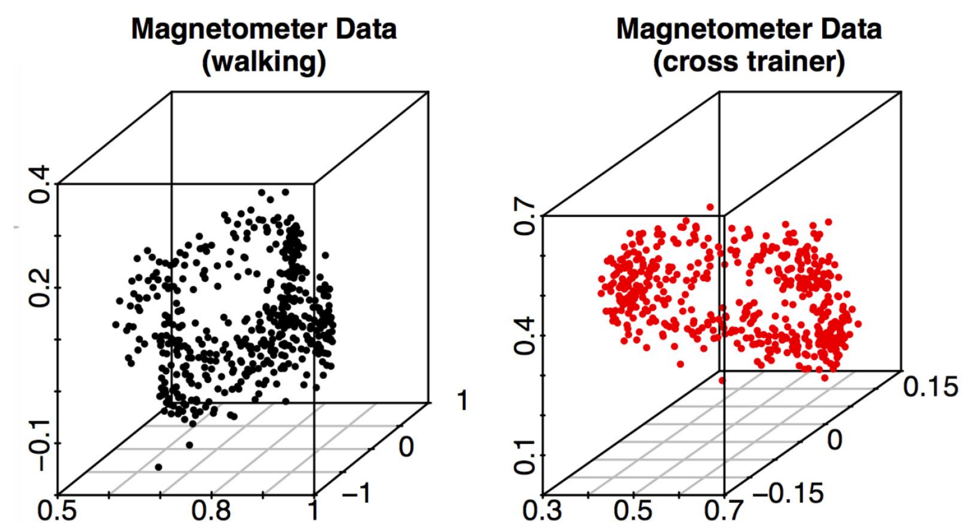
Introduction

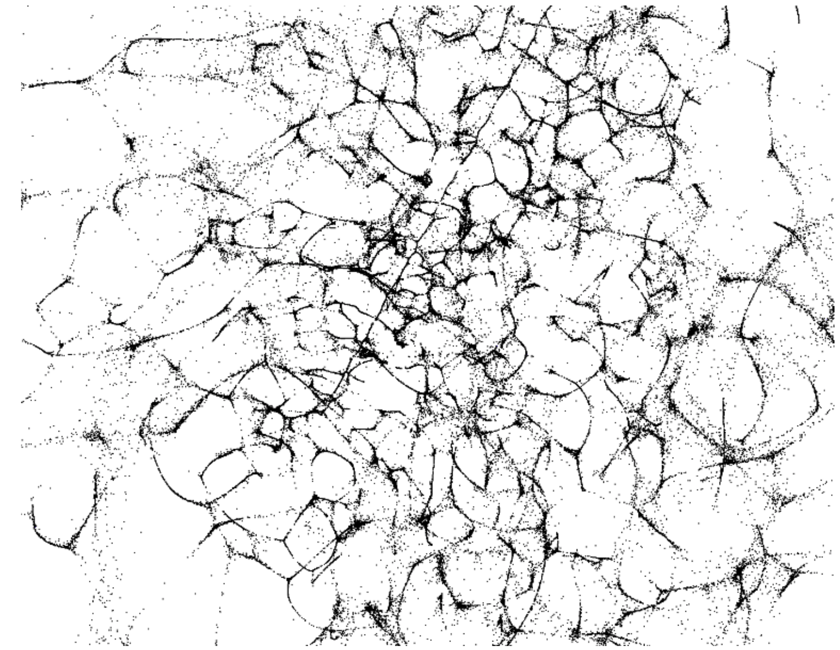


[Scanned 3D object]



[Shape database]

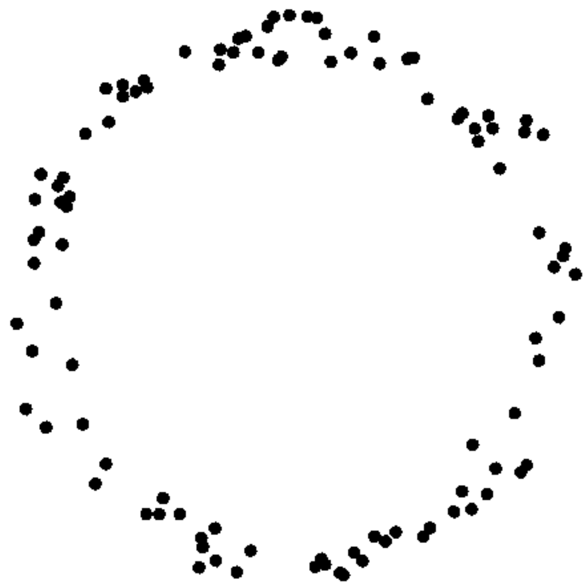


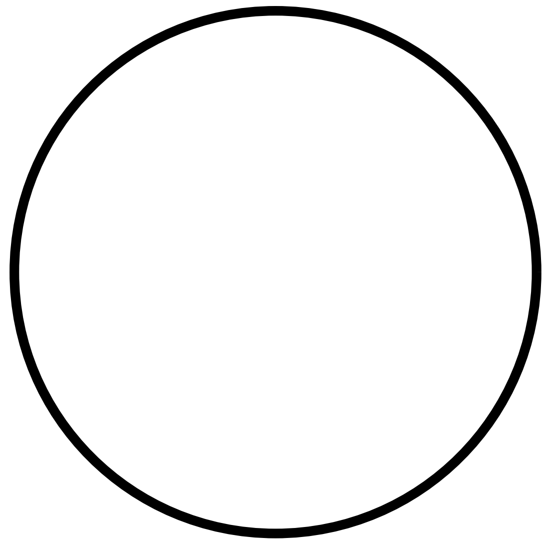


[Galaxies data]

- Data often come as (sampling of) metric spaces or sets/spaces endowed with a similarity measure with, possibly complex, topological/geometric structure.
- Data carrying geometric information are becoming high dimensional.
- Topological Data Analysis (TDA):
  - infer relevant topological and geometric features of these spaces.
  - take advantage of topol./geom. information for further processing of data (classification, recognition, learning, clustering, parametrization...).

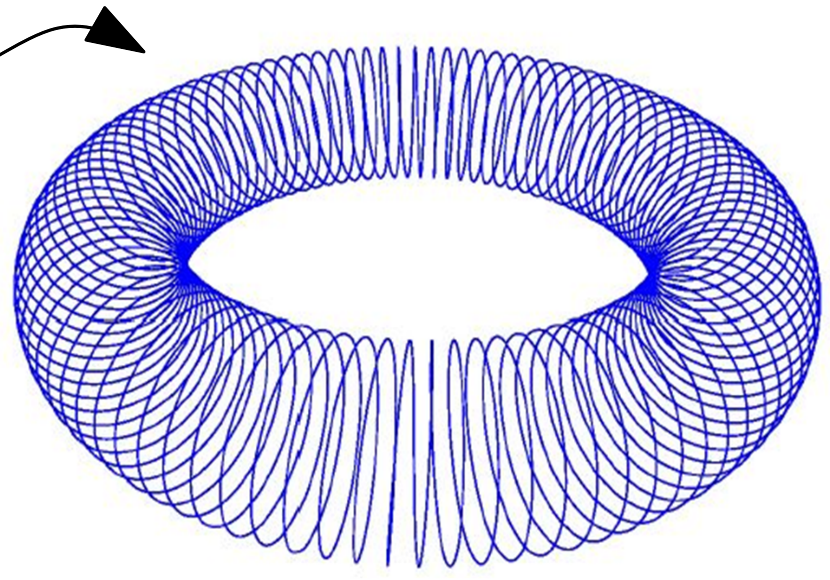
Challenges and goals



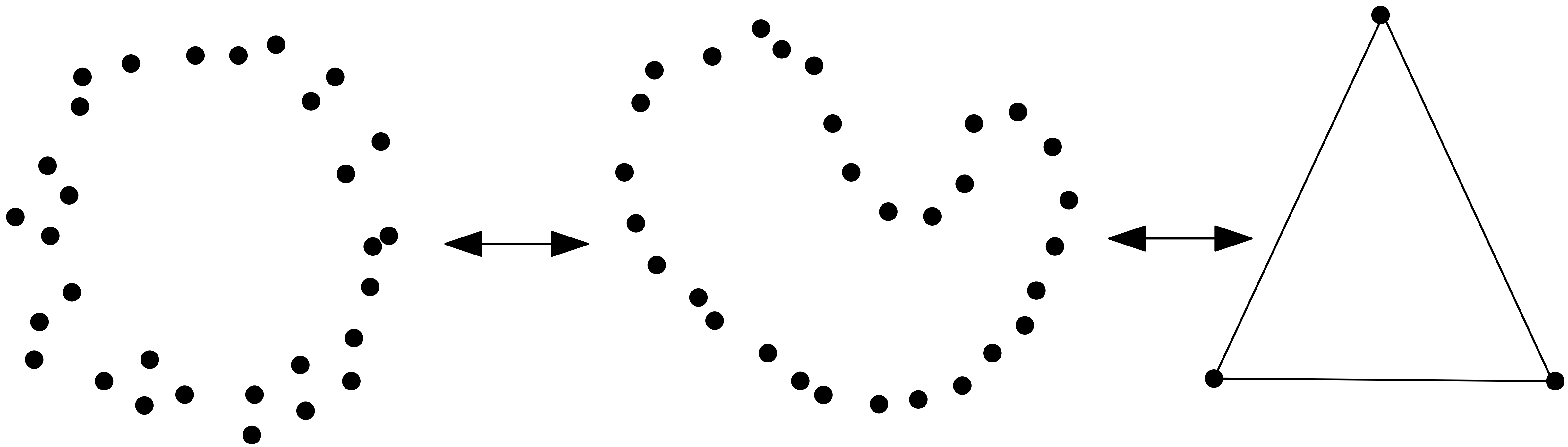


Problem(s):
- how to visualize the topological structure of data?
- how to compare topological properties (invariants) of close shapes/data sets?

- Challenges and goals:
  $\rightarrow$ no direct access to topological/geometric information: need of intermediate constructions (simplicial complexes);
  $\rightarrow$ distinguish topological "signal" from noise;
  $\rightarrow$ topological information may be multiscale;
  $\rightarrow$ statistical analysis of topological information.

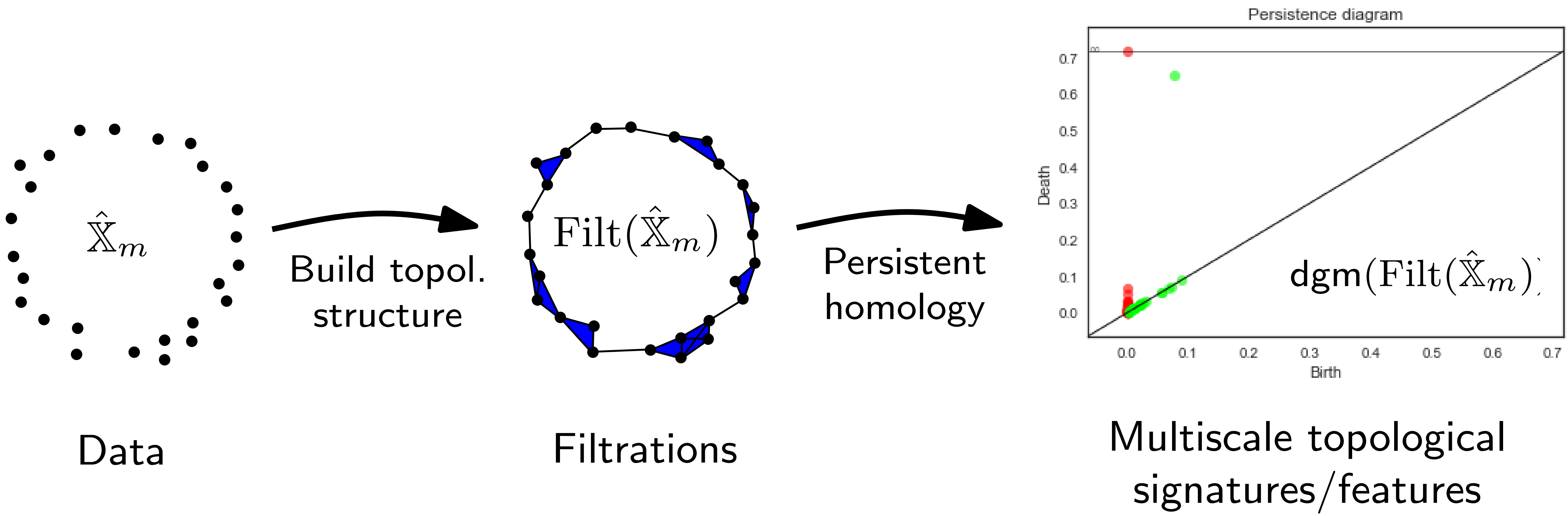


Why is topology interesting for data analysis?



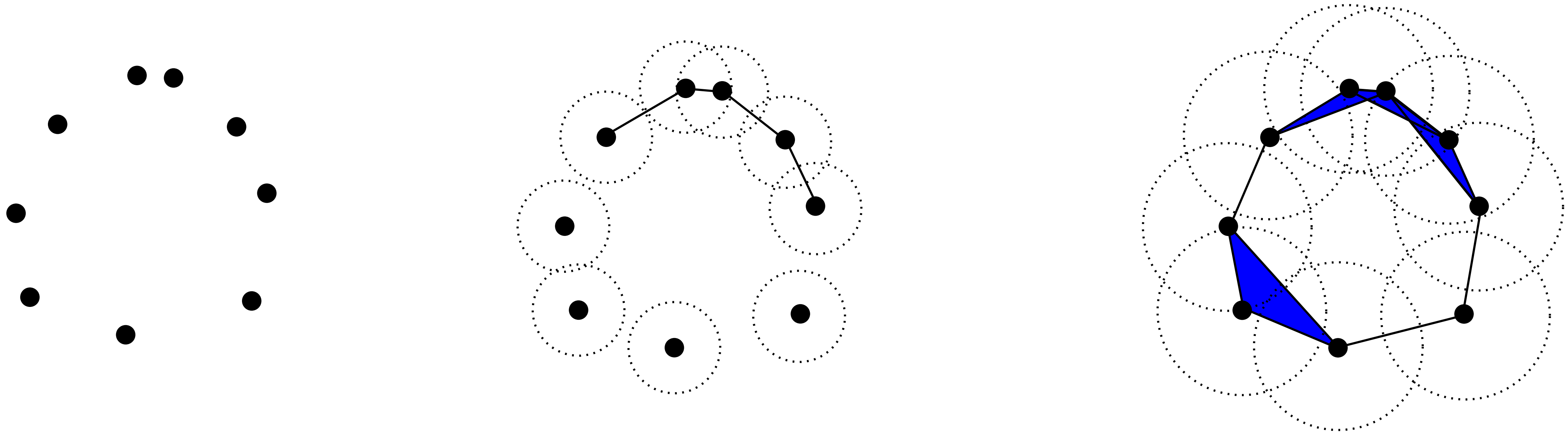
- **Coordinate invariance:** topological features/invariants do not rely on any coordinate system. ⇒ no need to have data with coordinate or to embed data in spaces with coordinates... But the metric (distance/similarity between data points) is important.
- **Deformation invariance:** topological features are invariant under homeomorphism.
- **Compressed representation:** Topology offer a set of tools to summarize and represent the data in compact ways while preserving its global topological structure.

The TDA pipeline



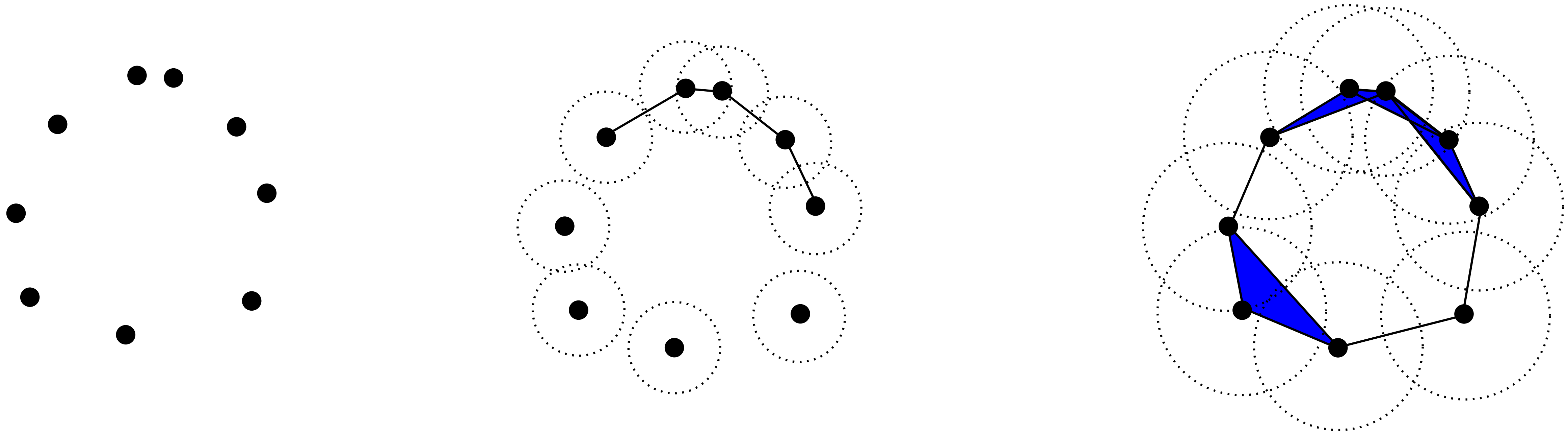
- Build a geometric filtered simplicial complex on top of $\widehat{\mathbb{X}}_m$ → multiscale topol. structure.
- Compute the persistent homology of the complex → multiscale topol. signature.
- Compare the signatures of “close” data sets → robustness and stability results.
- Statistical properties of signatures (connections with stability properties); use of topological information for further processing (e.g. Machine Learning).

Filtrations of simplicial complexes



A filtered simplicial complex $\mathbb{S}$ built on top of a set $\mathbb{X}$ is a family $(\mathbb{S}_a \mid a \in \mathbf{R})$ of subcomplexes of some fixed simplicial complex $\overline{\mathbb{S}}$ with vertex set $\mathbb{X}$ s. t. $\mathbb{S}_a \subseteq \mathbb{S}_b$ for any $a \leq b$.

Filtrations of simplicial complexes



A filtered simplicial complex $\mathbb{S}$ built on top of a set $\mathbb{X}$ is a family $(\mathbb{S}_a \mid a \in \mathbf{R})$ of subcomplexes of some fixed simplicial complex $\overline{\mathbb{S}}$ with vertex set $\mathbb{X}$ s. t. $\mathbb{S}_a \subseteq \mathbb{S}_b$ for any $a \leq b$.

Examples: Let $(\mathbb{X}, d_{\mathbb{X}})$ be a metric space.

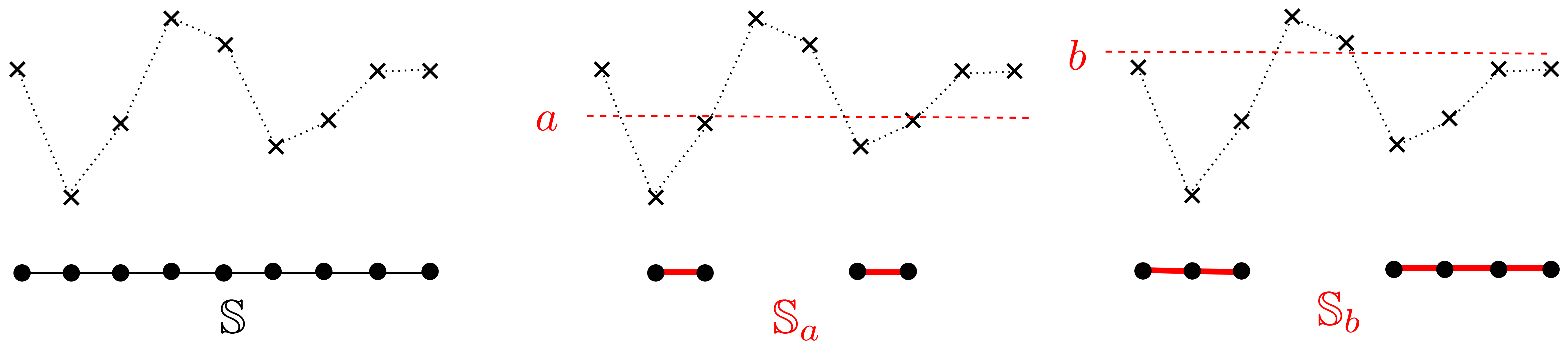
- The Vietoris-Rips filtration is the filtered simplicial complexe defined by: for $a \in \mathbf{R}$,

$$[x_0, x_1, \cdots, x_k] \in \mathrm{Rips}(\mathbb{X}, a) \Leftrightarrow d_{\mathbb{X}}(x_i, x_j) \leq a, \quad \text{for all } i, j.$$

- Čech complex: $\check{\mathrm{C}}\mathrm{ech}(\mathbb{X}, a)$ is the complex with vertex set $\mathbb{X}$ s.t.

$$[x_0, x_1, \cdots, x_k] \in \check{\mathrm{C}}\mathrm{ech}(\mathbb{X}, a) \Leftrightarrow \cap_{i=0}^{k} B(x_i, a) \neq \emptyset$$

Filtrations of simplicial complexes



A filtered simplicial complex $\mathbb{S}$ built on top of a set $\mathbb{X}$ is a family $(\mathbb{S}_a \mid a \in \mathbf{R})$ of subcomplexes of some fixed simplicial complex $\overline{\mathbb{S}}$ with vertex set $\mathbb{X}$ s. t. $\mathbb{S}_a \subseteq \mathbb{S}_b$ for any $a \leq b$.

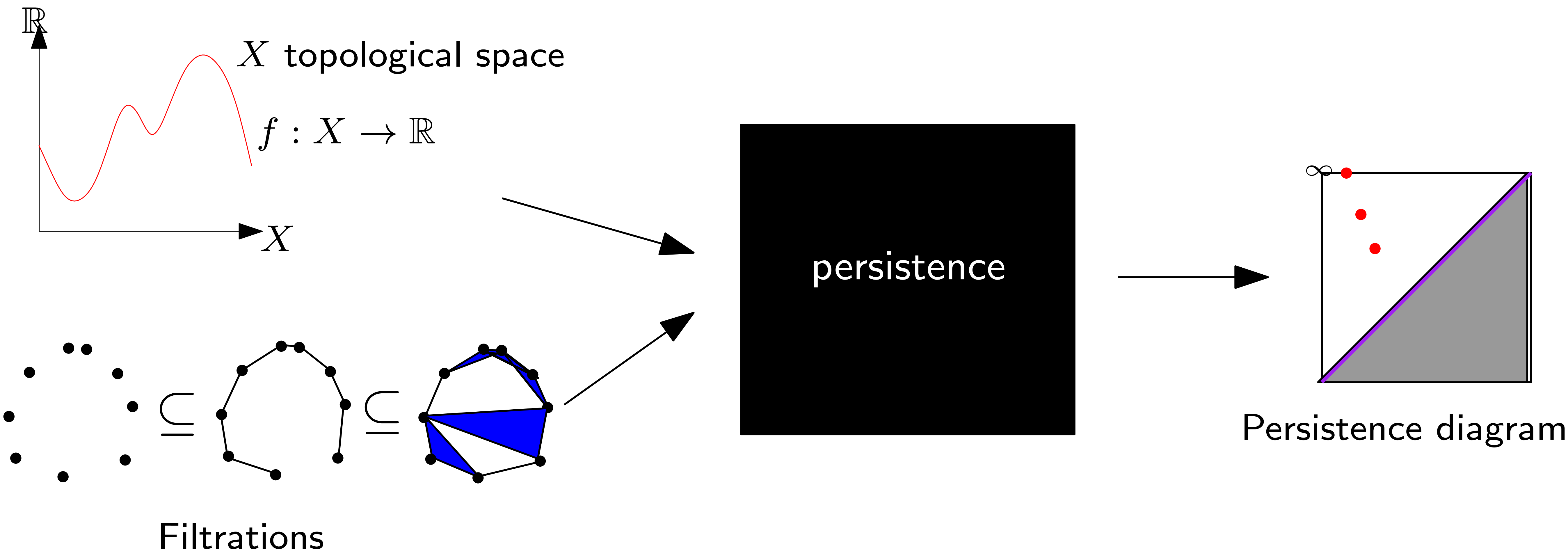
Examples:
Let $\mathbb{S}$ be a simplicial complex with vertex set $\mathbb{X}$ and let $f : \mathbb{X} \to \mathbb{R}$.

For $\sigma = [v_0, \cdots, v_k] \in \mathbb{S}$, define $f(\sigma) = \max\{f(v_i) : i = 0, \cdots, k\}$.

The sublevel set filtration of f is the family of subcomplexes

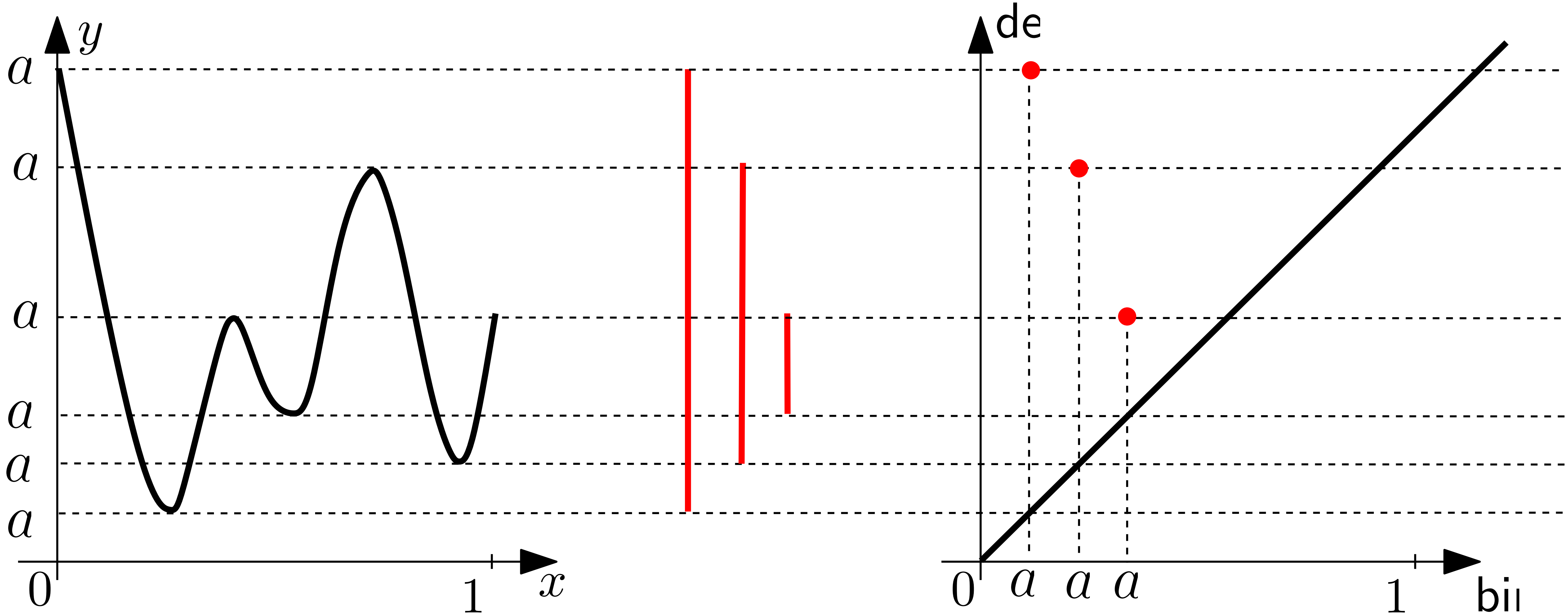
$$\mathbb{S}_a = \{\sigma \in \mathbb{S} : f(\sigma) \leq a\}, a \in \mathbb{R}.$$

Persistent homology



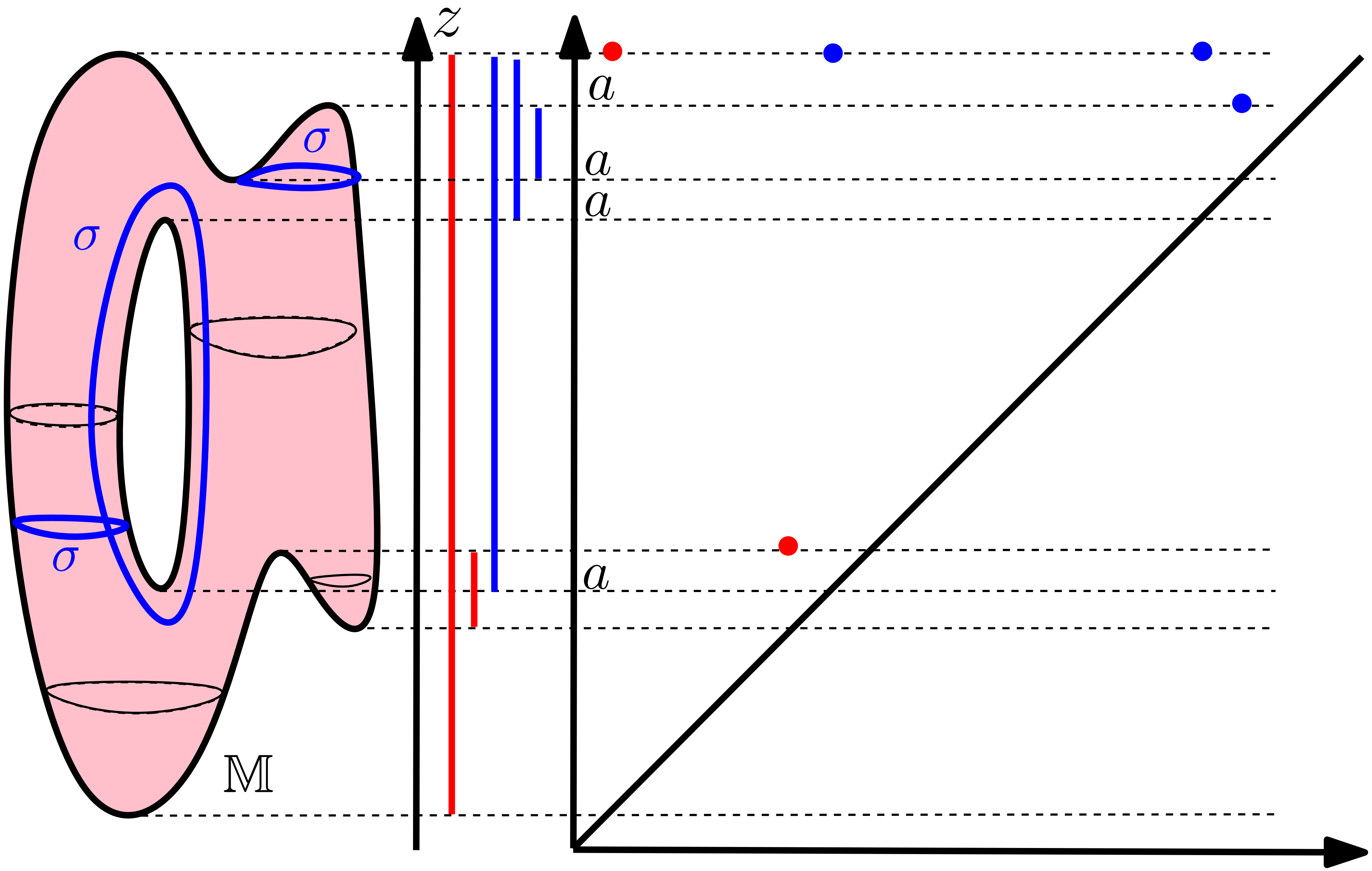
- A general mathematical framework to encode the evolution of the topology (homology) of families of nested spaces (filtered complex, sublevel sets,...).
- Formalized by H. Edelsbrunner (2002) et al and G. Carlsson et al (2005) - wide development during the last decade. Ideas tracing back to M. Morse (1940)!
- Multiscale topological information.
- Barcodes/persistence diagrams can be efficiently computed (e.g. Gudhi library!).
- Stability properties

Persistent homology for functions



Tracking and encoding the evolution of the connected components (0-dimensional homology) of the sublevel sets of a function

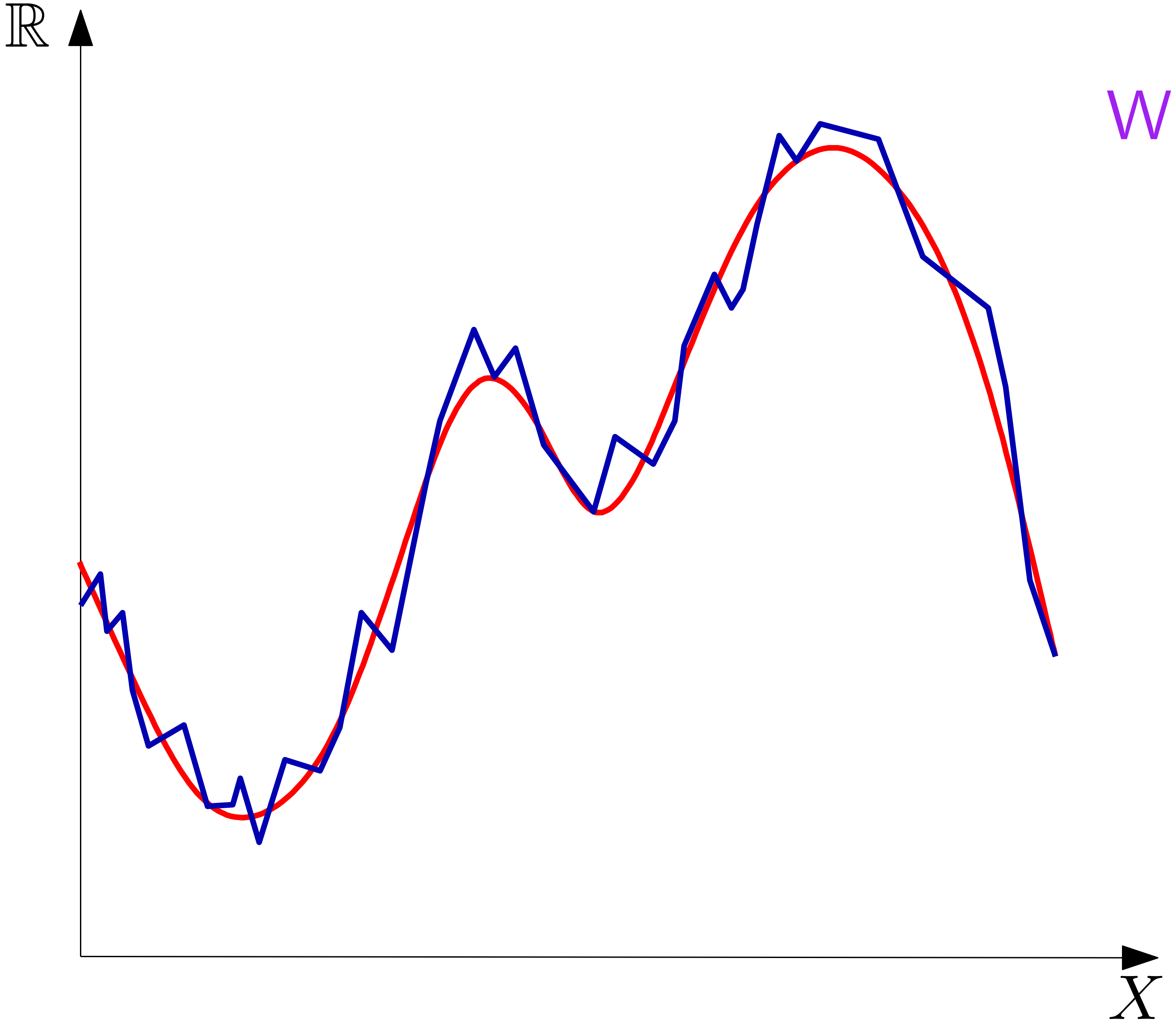
Persistent homology for functions



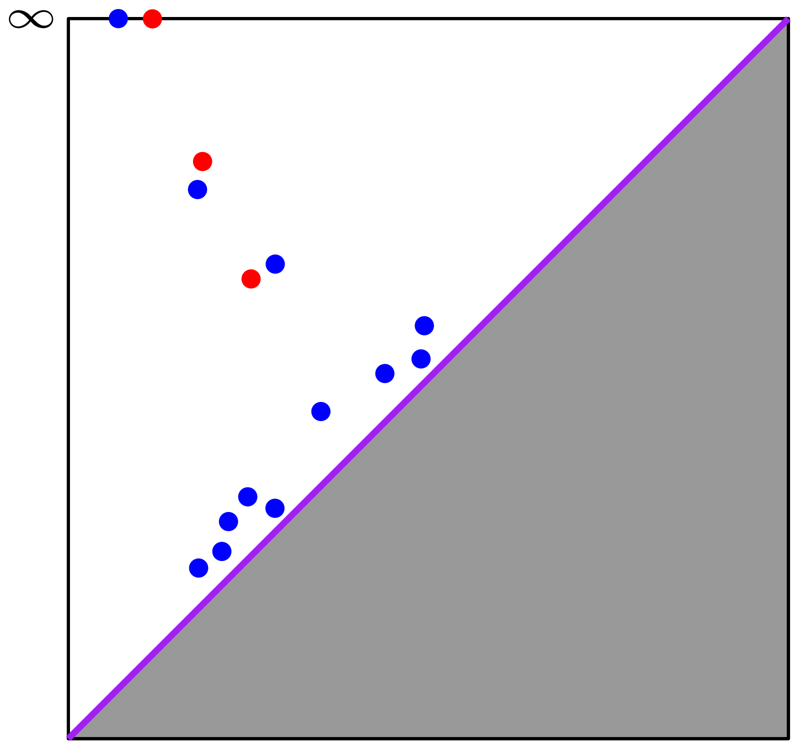
Tracking and encoding the evolution of the connected components (0-dimensional homology) and cycles (1-dimensional homology) of the sublevel sets.

Homology: an algebraic way to rigorously formalize the notion of k-dimensional cycles through a vector space (or a group), the homology group whose dimension is the number of "independent" cycles (the Betti number).

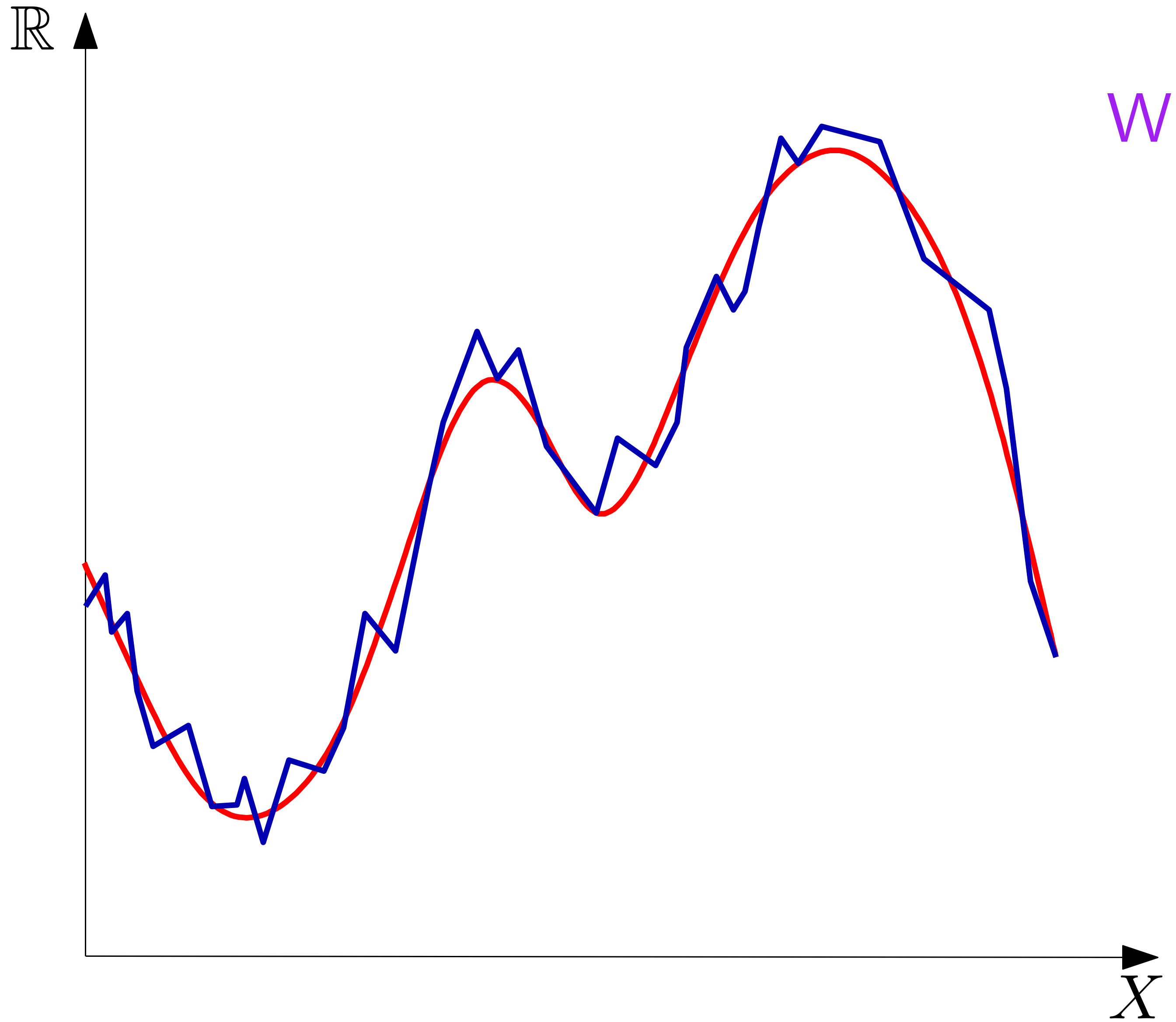
Stability properties



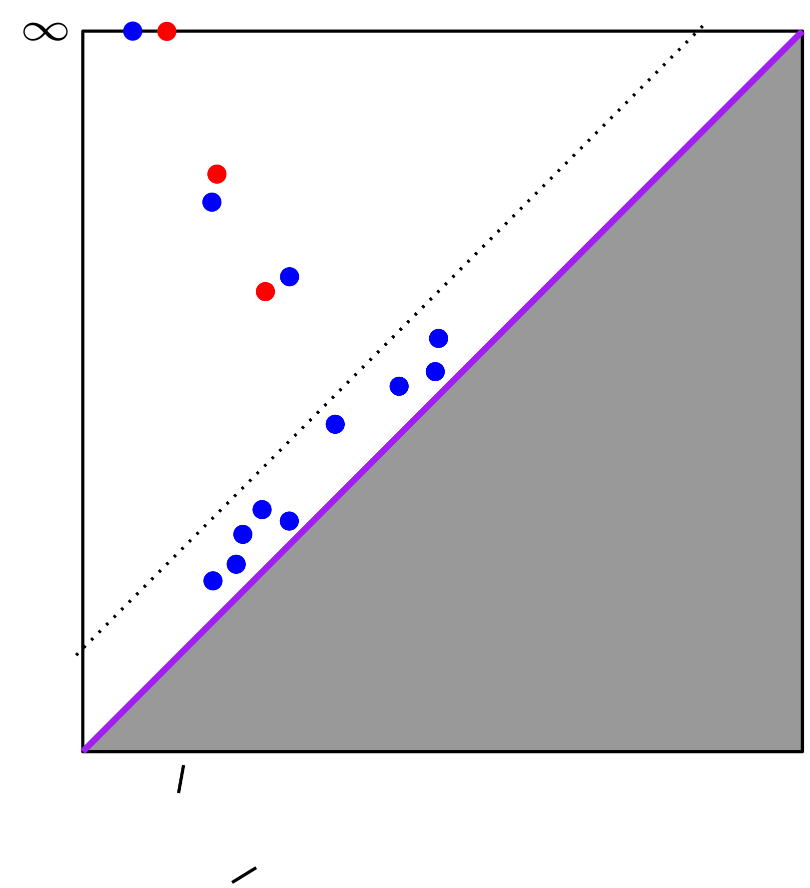
What if f is slightly perturbed?



Stability properties



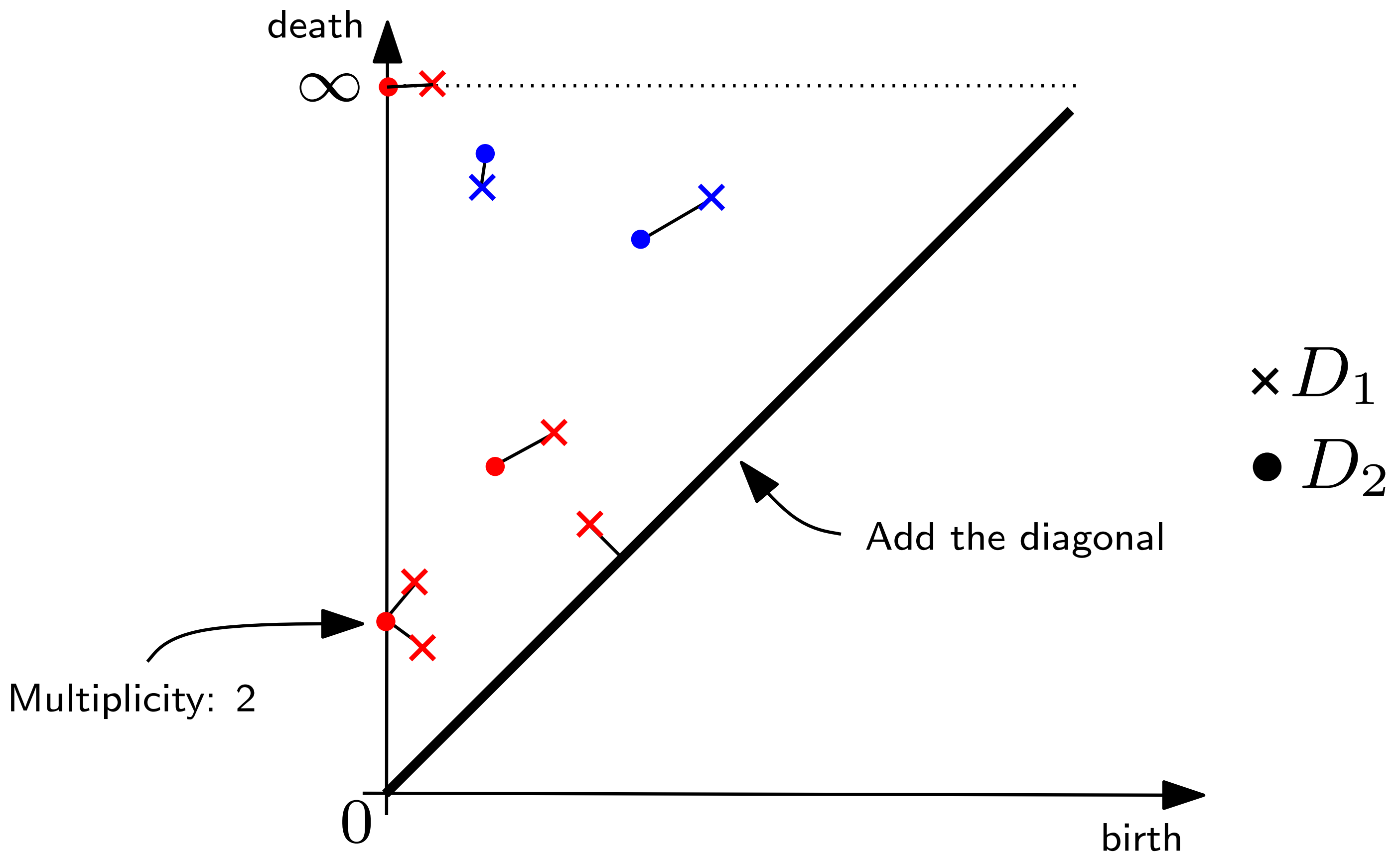
What if f is slightly perturbed?



Theorem (Stability):
For any *tame* functions $f, g : \mathbb{X} \to \mathbb{R}$, $d_B(\mathrm{D}_f, \mathrm{D}_g) \leq \|f - g\|_\infty$.

[Cohen-Steiner, Edelsbrunner, Harer 05], [C., Cohen-Steiner, Glisse, Guibas, Oudot - SoCG 09], [C., de Silva, Glisse, Oudot 12]

Comparing persistence diagrams



The bottleneck distance between two diagrams D_1 and D_2 is

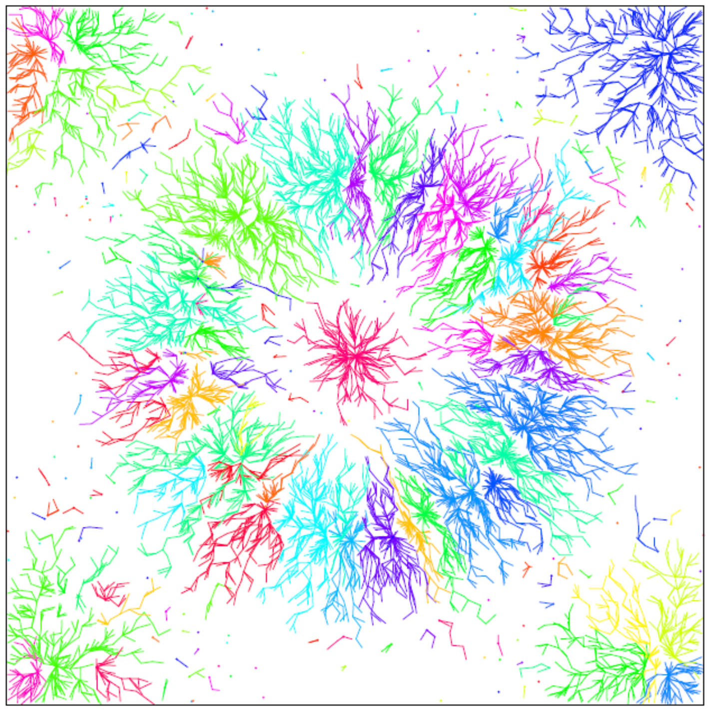
$$d_B(D_1, D_2) = \inf_{\gamma \in \Gamma} \sup_{p \in D_1} \|p - \gamma(p)\|_\infty$$

where Γ is the set of all the bijections between D_1 and D_2 and $\|p - q\|_\infty = \max(|x_p - x_q|, |y_p - y_q|)$.

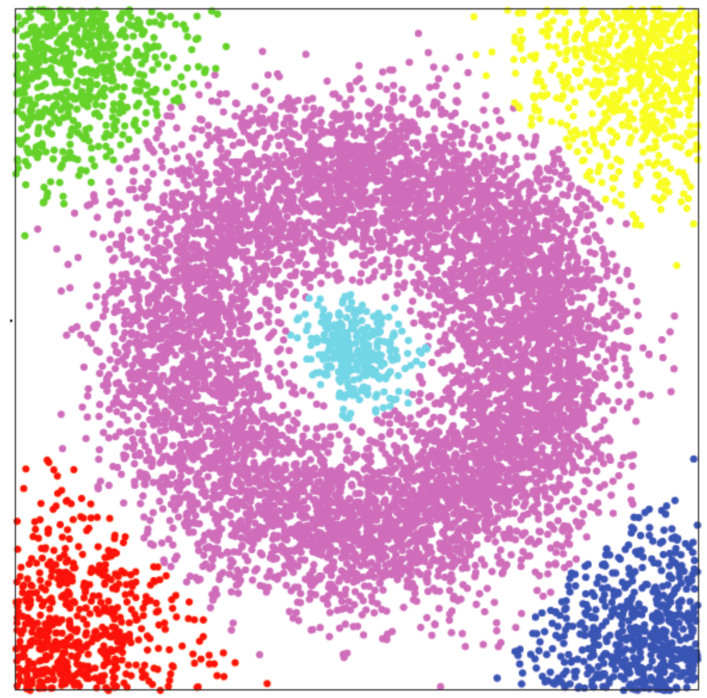
$\rightarrow$ Persistence diagrams provide easy to compare topological signatures.

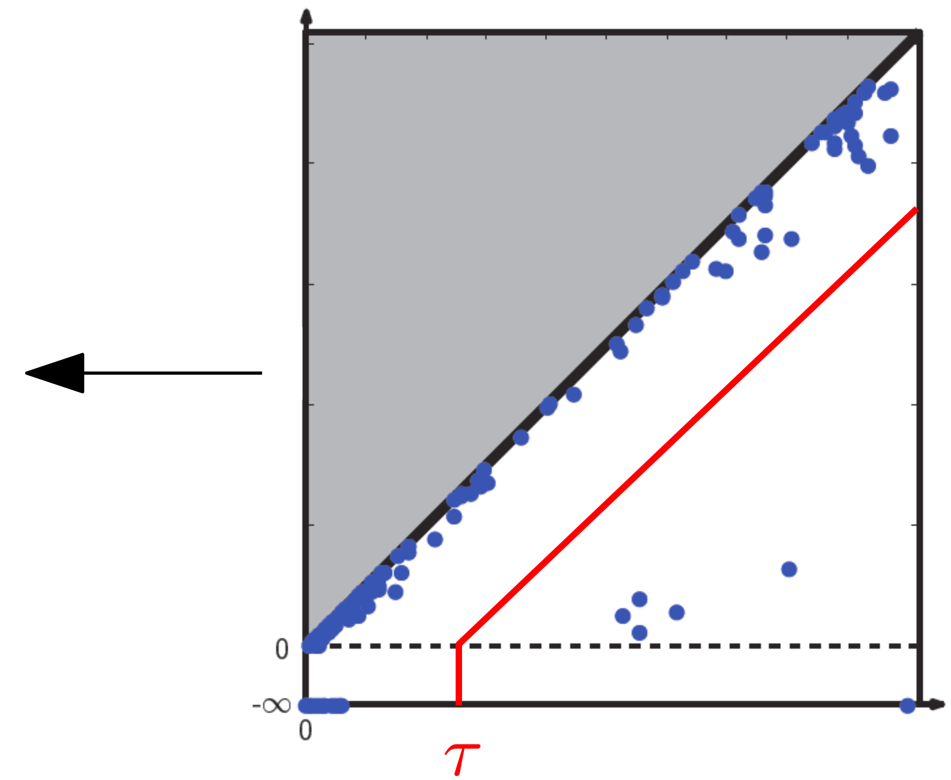
Some examples of applications

- Persistence-based clustering [C.,Guibas,Oudot,Skraba - J. ACM 2013]

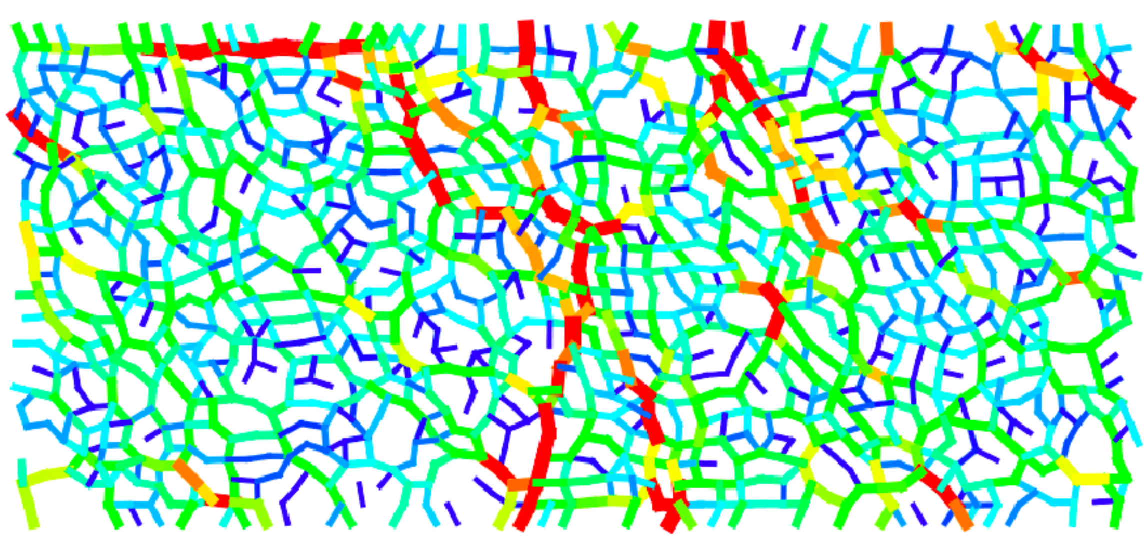


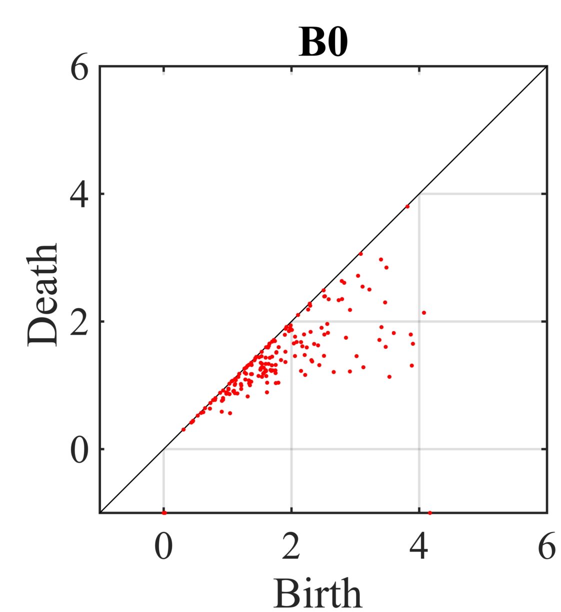
$\tau = 0$





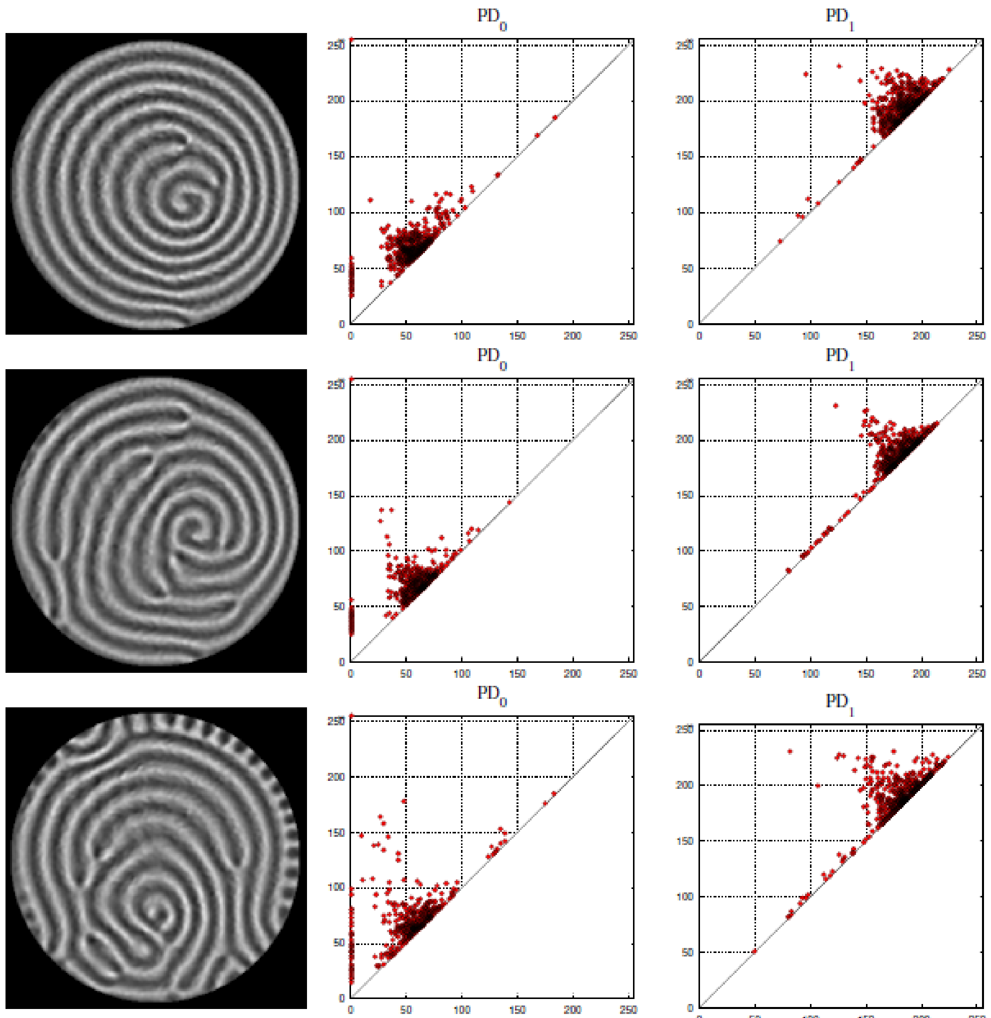
- Analysis of force fields in granular media [Kramar, Mischaikow et al]

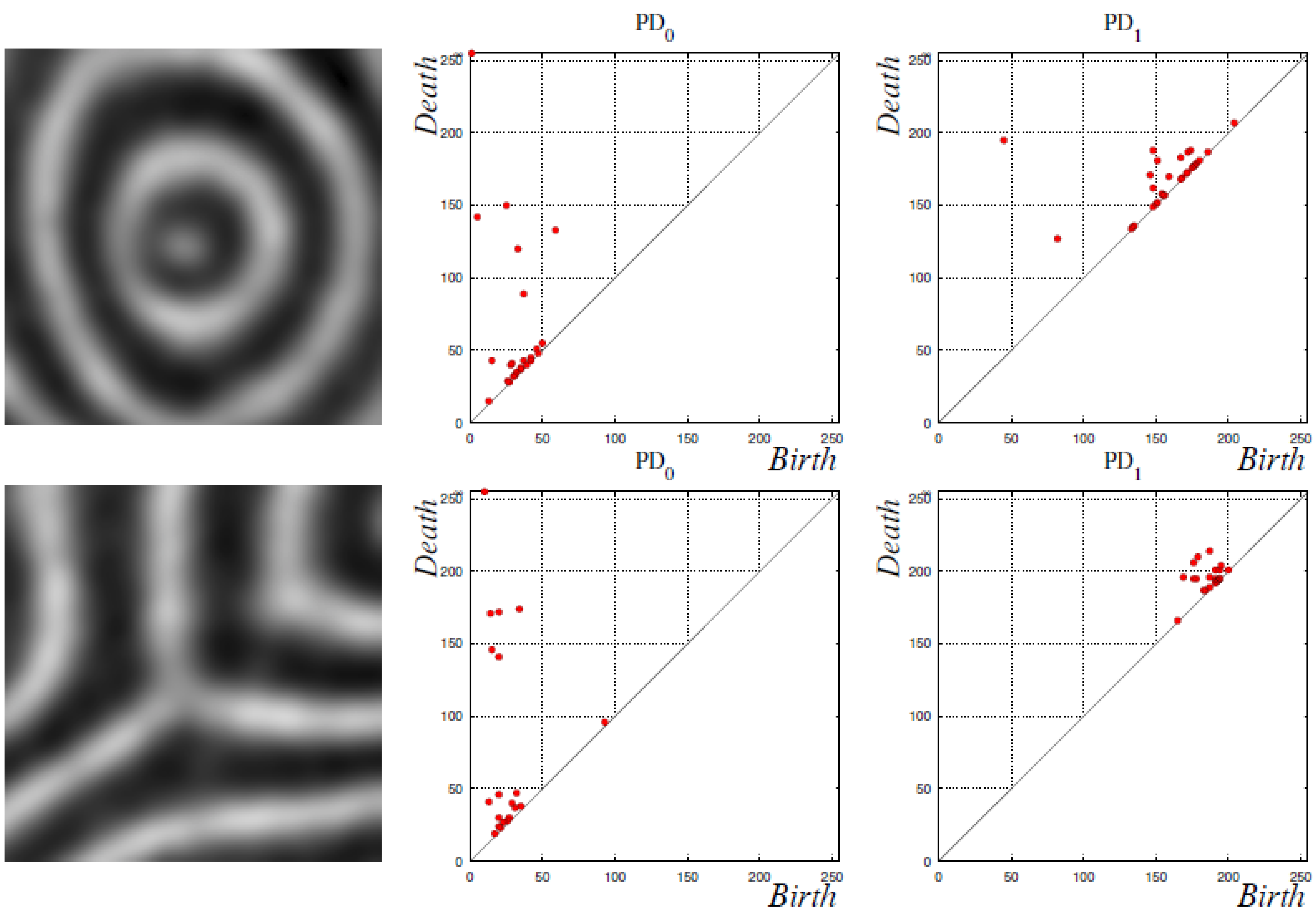




Some examples of applications

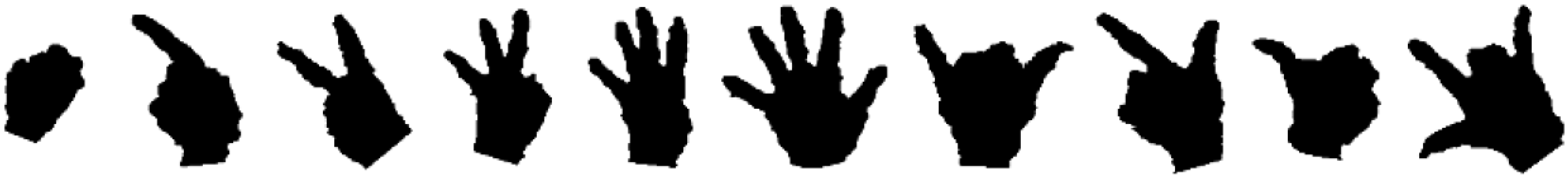
- Pattern analysis in fluid dynamics [Kramar, Mischaikow et al]

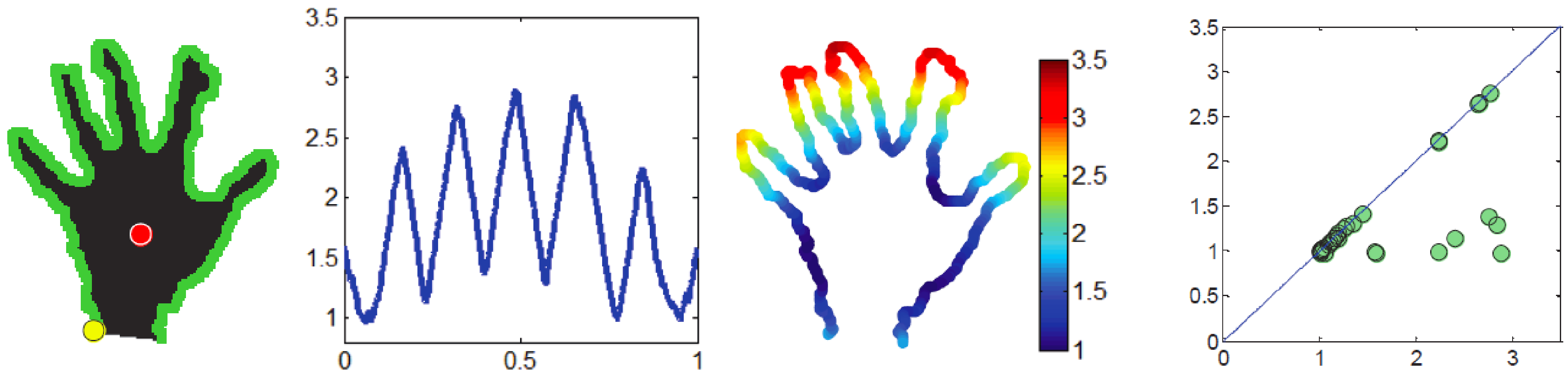




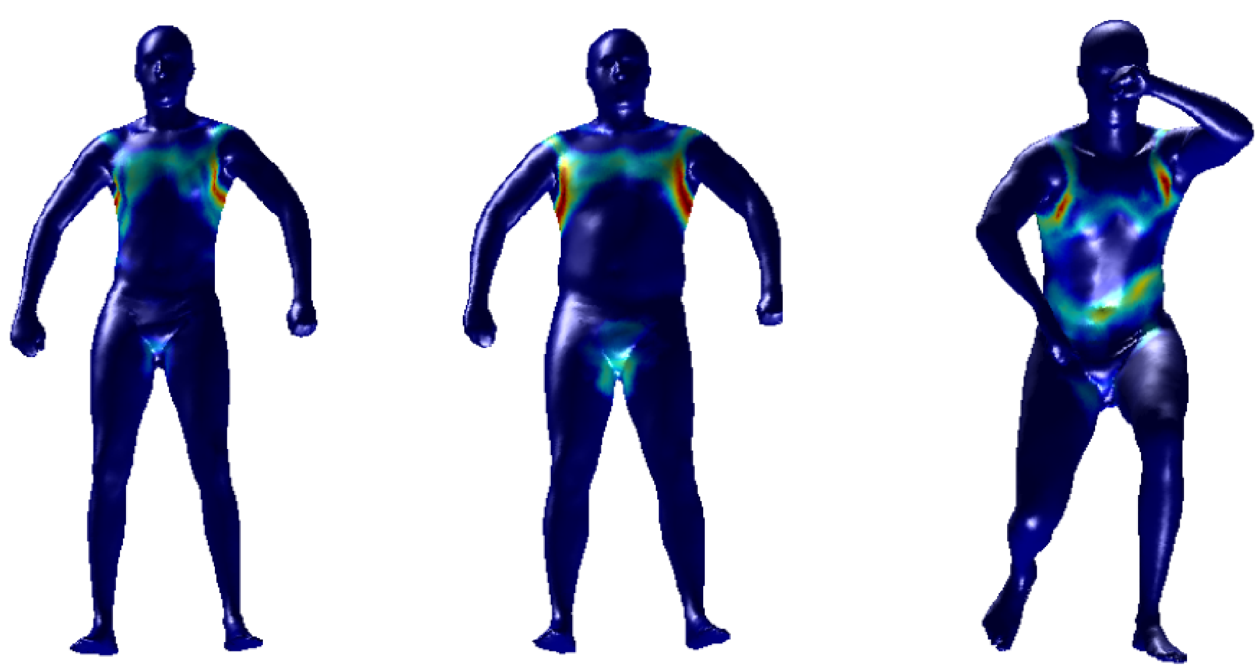
Some examples of applications

- Hand gesture recognition [Li, Ovsjanikov, C. - CVPR'14]

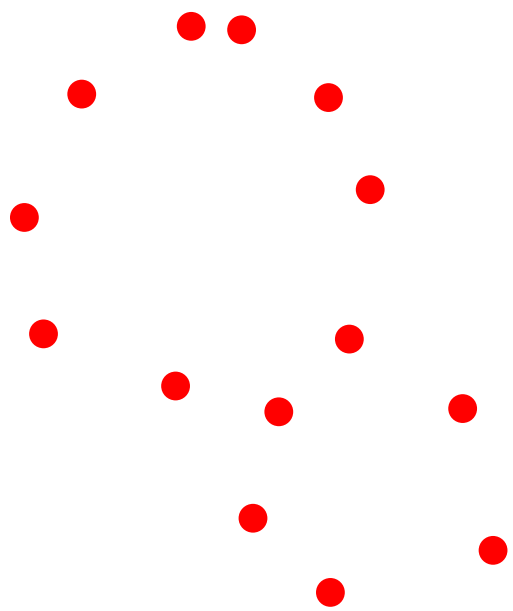




- Persistence-based pooling for shape recognition [Bonis, Ovsjanikov, Oudot, C. 2016]

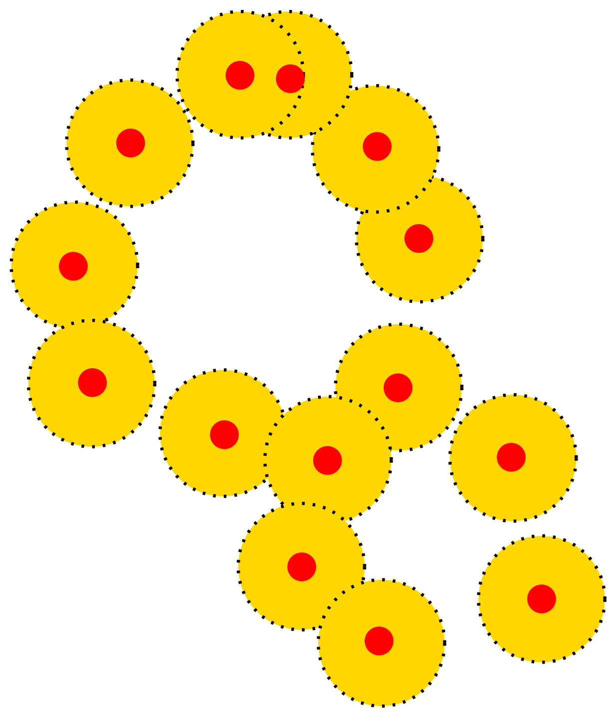


Persistent homology for point cloud data



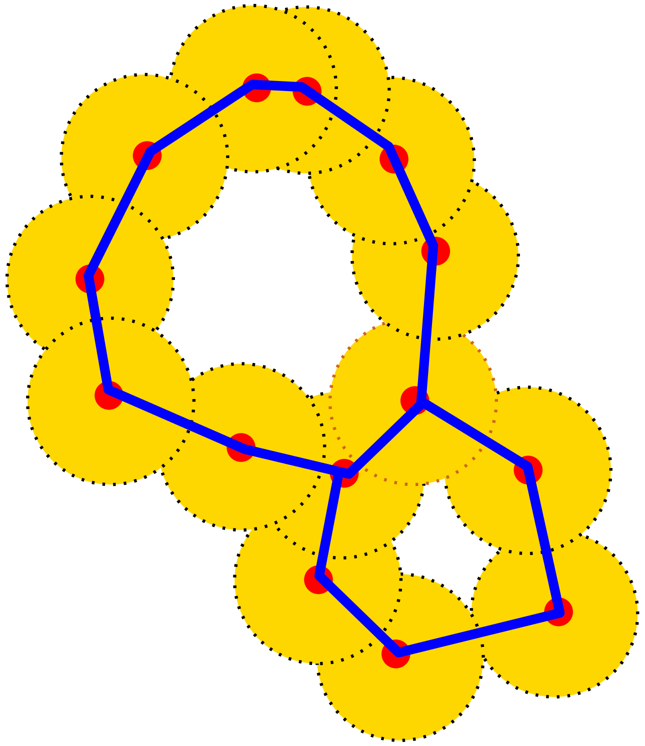
- Filtrations allow to construct “shapes” representing the data in a multiscale way.
- Persistent homology: encode the evolution of the topology across the scales $\rightarrow$ multi-scale topological signatures.

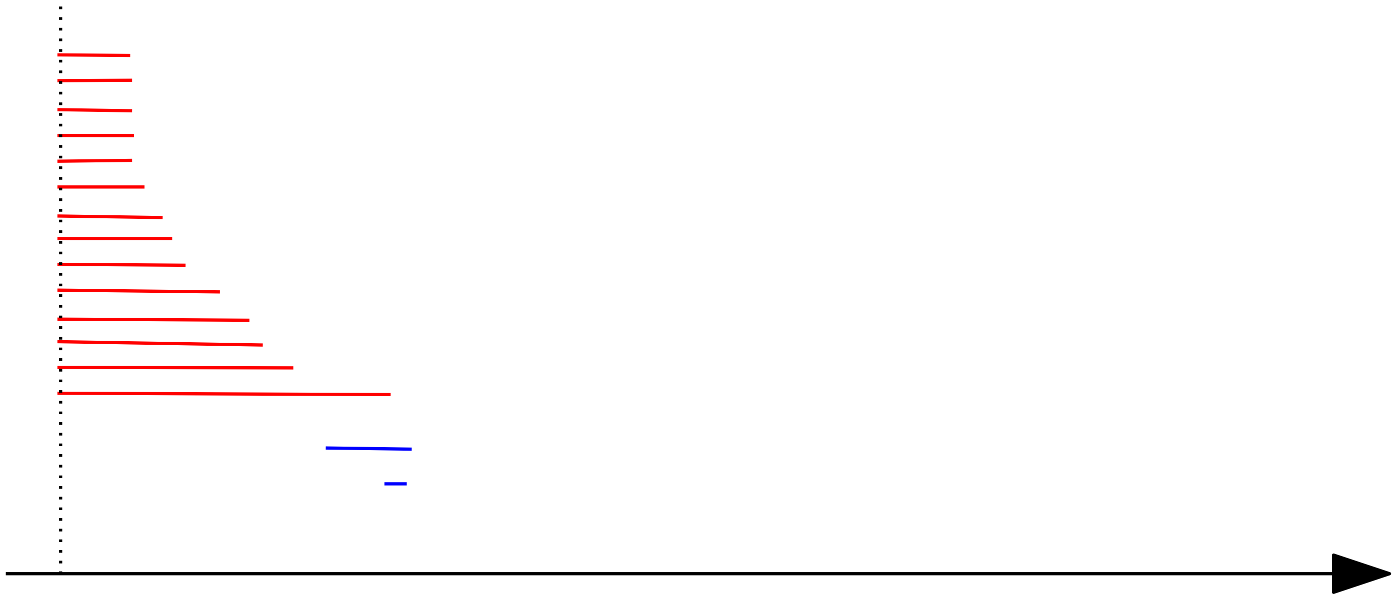
Persistent homology for point cloud data



- Filtrations allow to construct “shapes” representing the data in a multiscale way.
- Persistent homology: encode the evolution of the topology across the scales $\rightarrow$ multi-scale topological signatures.

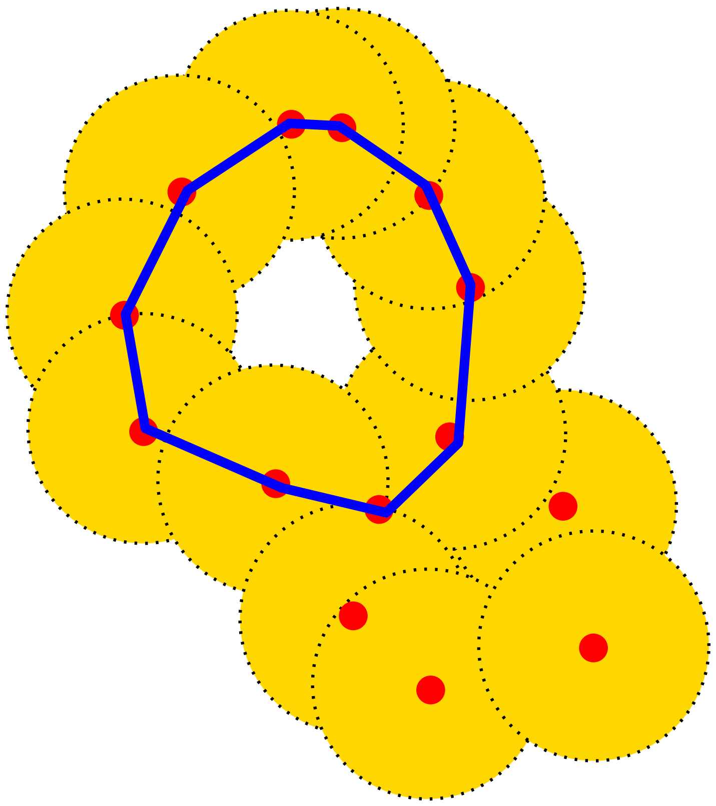
Persistent homology for point cloud data

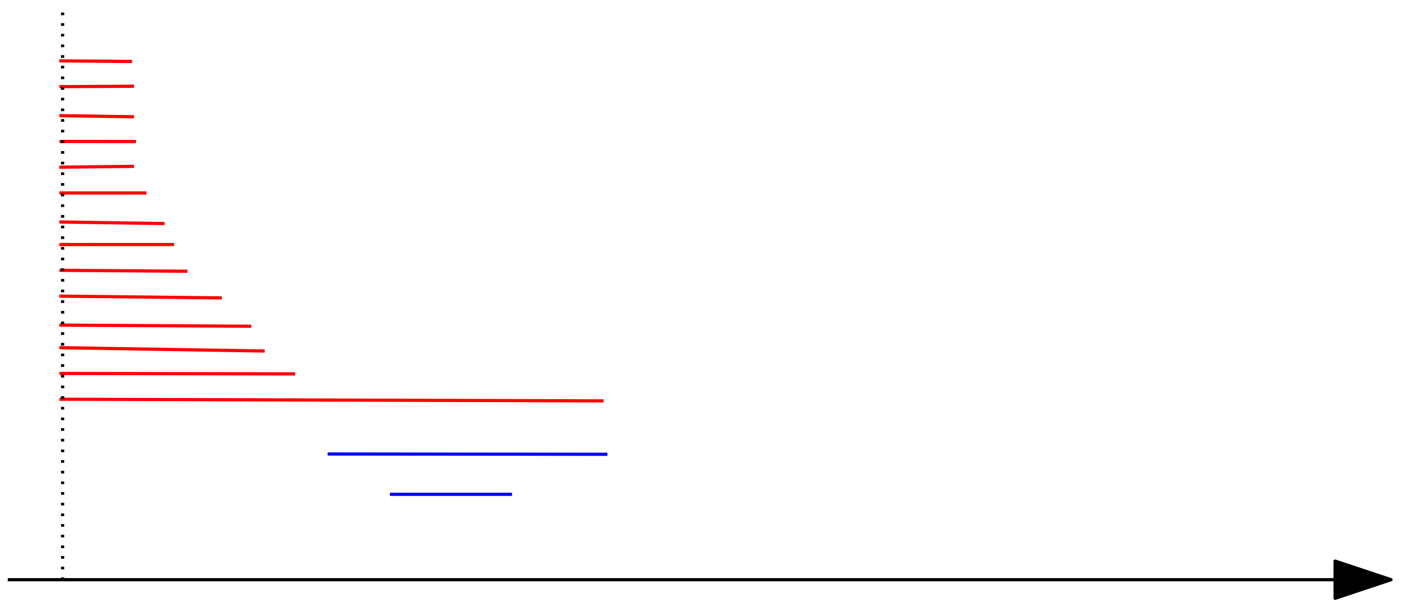




- Filtrations allow to construct "shapes" representing the data in a multiscale way.
- Persistent homology: encode the evolution of the topology across the scales $\rightarrow$ multi-scale topological signatures.

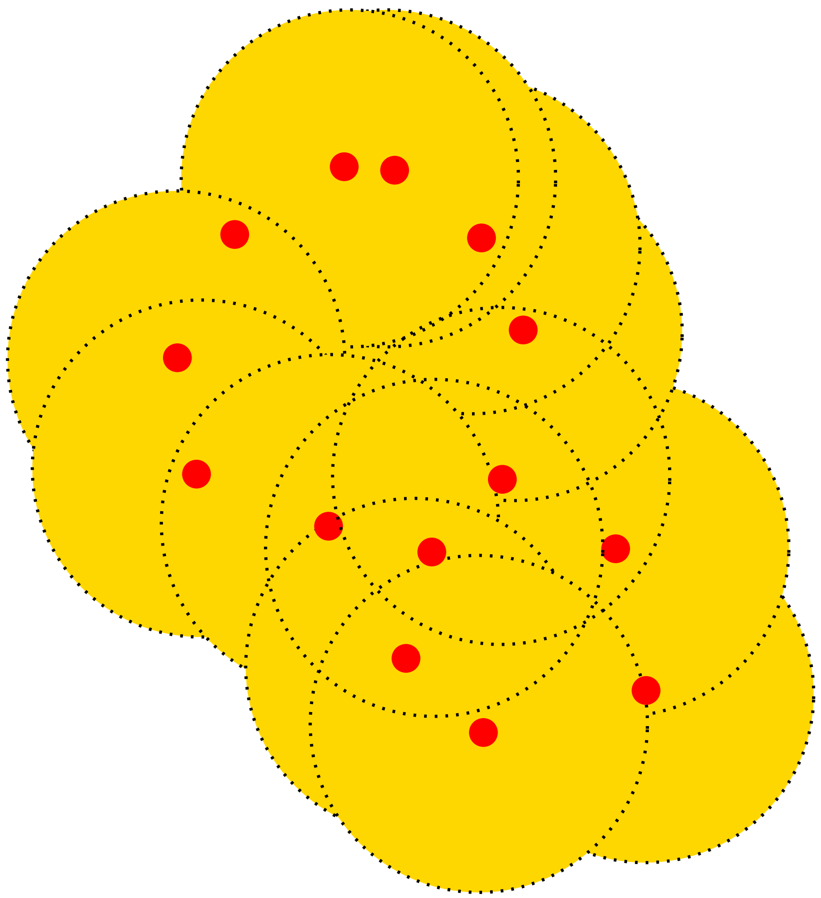
Persistent homology for point cloud data

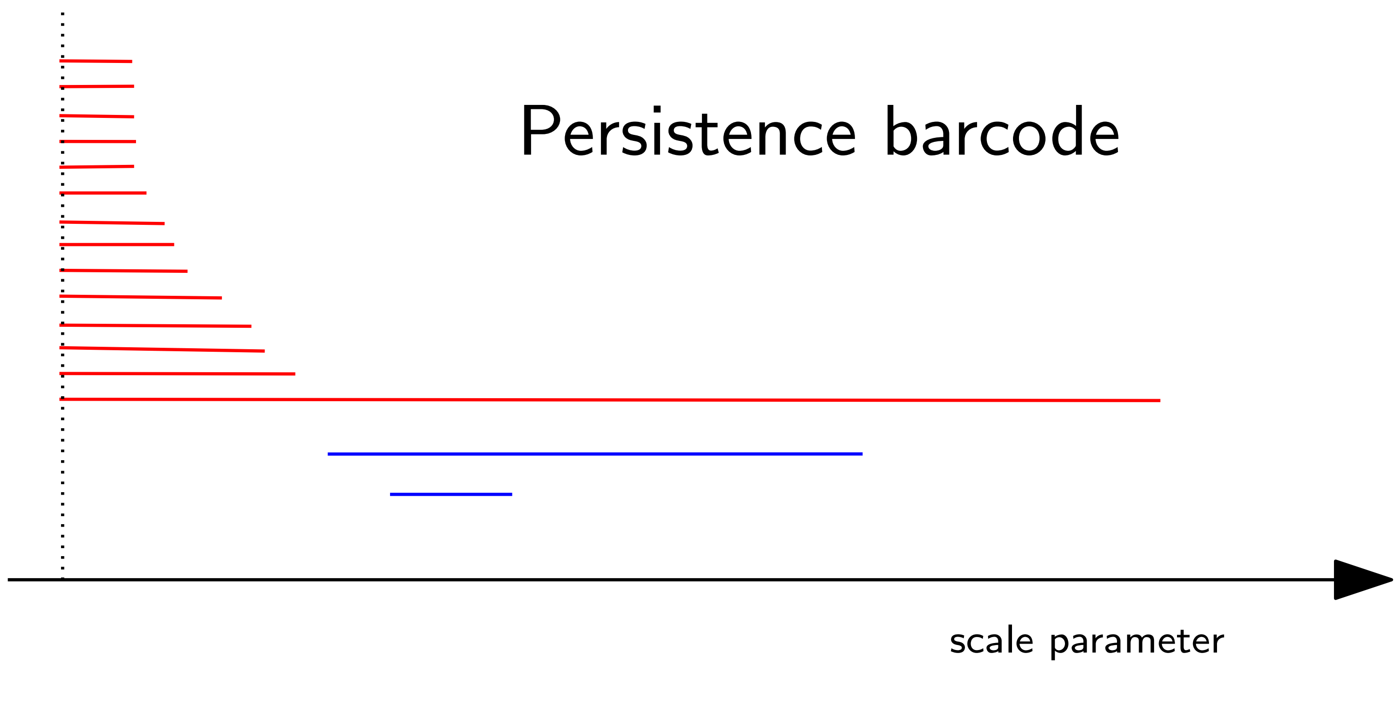


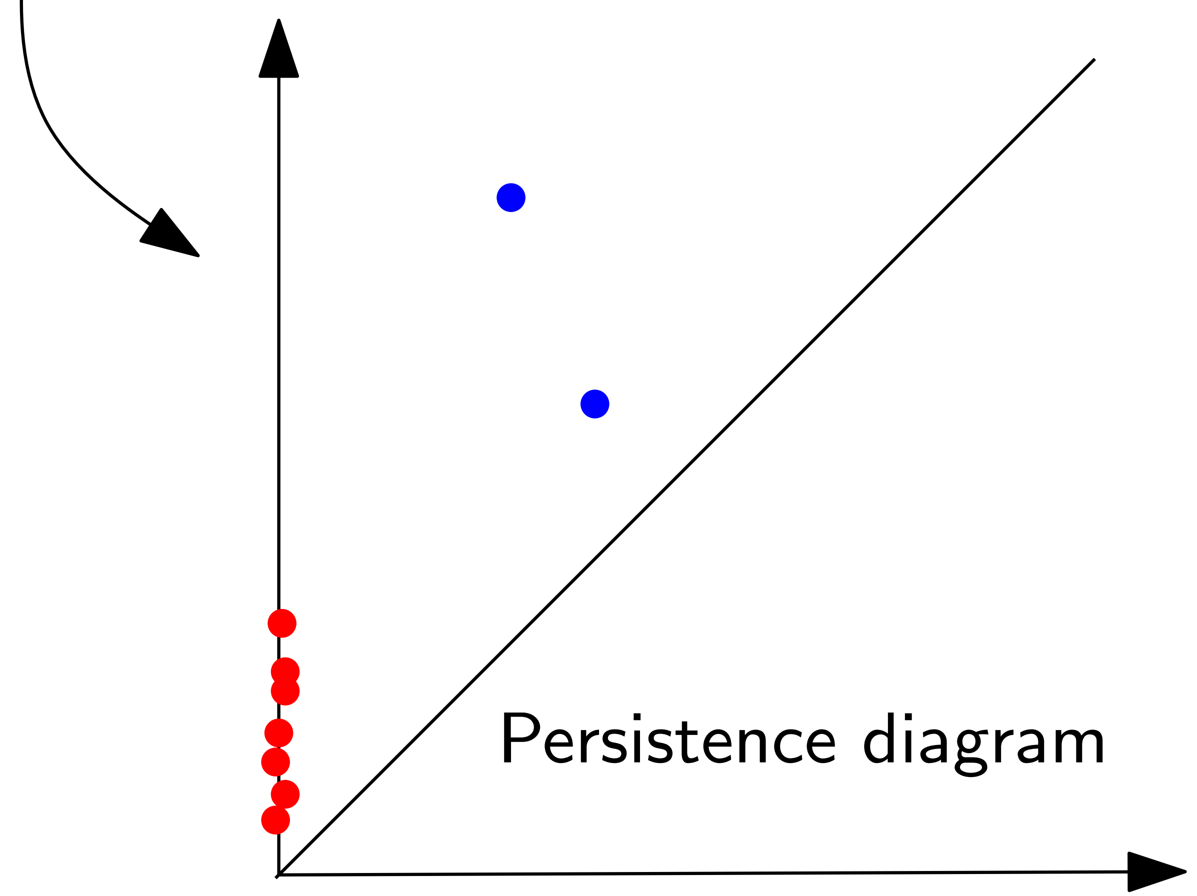


- Filtrations allow to construct "shapes" representing the data in a multiscale way.
- Persistent homology: encode the evolution of the topology across the scales $\rightarrow$ multi-scale topological signatures.

Persistent homology for point cloud data







- Filtrations allow to construct "shapes" representing the data in a multiscale way.
- Persistent homology: encode the evolution of the topology across the scales $\rightarrow$ multi-scale topological signatures.

Stability properties

"Stability theorem": Close spaces/data sets have close persistence diagrams!

[C., de Silva, Oudot - Geom. Dedicata 2013].

If $\mathbb{X}$ and $\mathbb{Y}$ are pre-compact metric spaces, then

$$\mathrm{d_b}(\mathsf{dgm}(\mathrm{Filt}(\mathbb{X})), \mathsf{dgm}(\mathrm{Filt}(\mathbb{Y}))) \leq d_{GH}(\mathbb{X}, \mathbb{Y}).$$

Bottleneck distance

Here Filt can be $\mathbb{R}\mathrm{ips}$, $\check{\mathrm{C}}\mathrm{ech}$, etc...

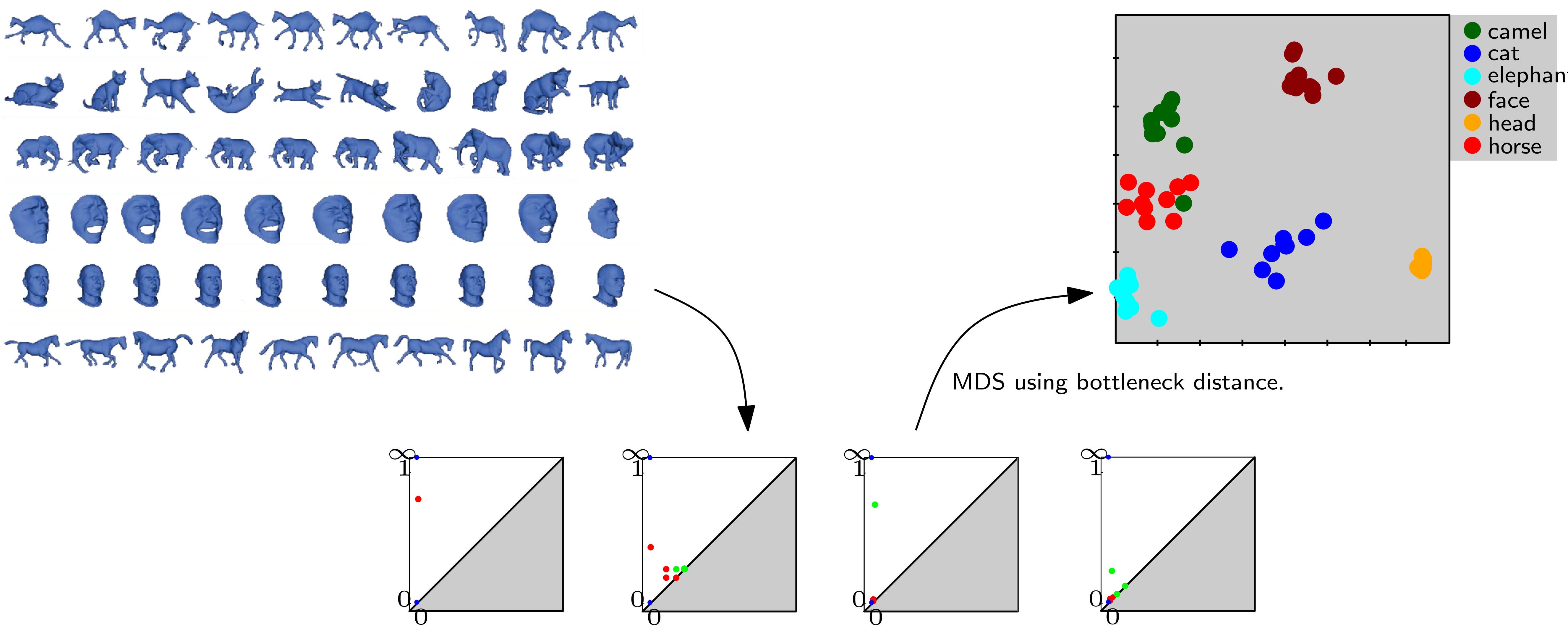
Gromov-Hausdorff distance

$$d_{GH}(\mathbb{X}, \mathbb{Y}) := \inf_{\mathbb{Z}, \gamma_1, \gamma_2} d_H(\gamma_1(\mathbb{X}), \gamma_2(\mathbb{X}))$$

$\mathbb{Z}$ metric space, $\gamma_1 : \mathbb{X} \to \mathbb{Z}$ and $\gamma_2 : \mathbb{Y} \to \mathbb{Z}$ isometric embeddings.

Application: non rigid shape classification

[C., Cohen-Steiner, Guibas, Mémoli, Oudot - SGP '09]



- Non rigid shapes in a same class are almost isometric, but computing Gromov-Hausdorff distance between shapes is extremely expensive.
- Compare diagrams of sampled shapes instead of shapes themselves.

The theory of persistence

Theory of persistence has been subject to intense research activities:

- **from the mathematical perspective:**

  - general algebraic framework (persistence modules) and general stability results.
  - extensions and generalizations of persistence (zig-zag persistence, multi-persistence, etc...)
  - Statistical analysis of persistence.

- **from the algorithmic and computational perspective:**

  - efficient algorithms to compute persistence and some of its variants.
  - efficient software libraries (in particular, Gudhi: https://project.inria.fr/gudhi/).

A whole machinery at the crossing of mathematics and computer science!

Some drawbacks and problems

If $\mathbb{X}$ and $\mathbb{Y}$ are pre-compact metric spaces, then

$$\mathrm{d_b}(\mathsf{dgm}(\mathrm{Rips}(\mathbb{X})), \mathsf{dgm}(\mathrm{Rips}(\mathbb{Y}))) \leq d_{GH}(\mathbb{X}, \mathbb{Y}).$$

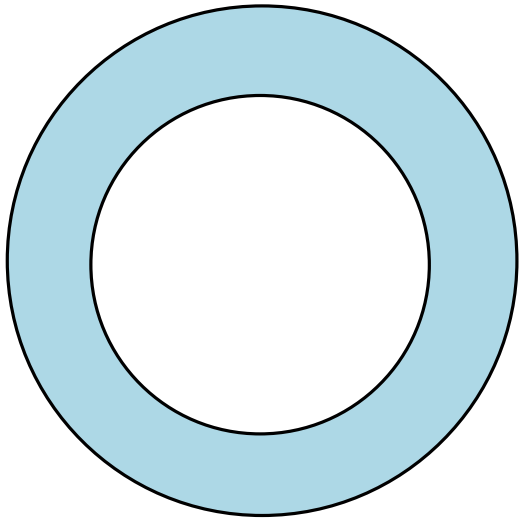
$\rightarrow$ Vietoris-Rips (or Cech,...) filtrations quickly become prohibitively large as the size of the data increases ($O(|\mathbb{X}|^d)$), making the computation of persistence of large data sets a real challenge.

$\rightarrow$ Persistence diagrams of Rips-Vietoris (and Cěch, witness,..) filtrations and Gromov-Hausdorff distance are very sensitive to noise and outliers.

$\rightarrow$ The space of persistence diagrams endowed with the bottleneck distance is highly non linear, processing persistence information for further data analysis and learning tasks is a challenge.

These issues have raised an intense research activity during the last few years!

Statistical setting

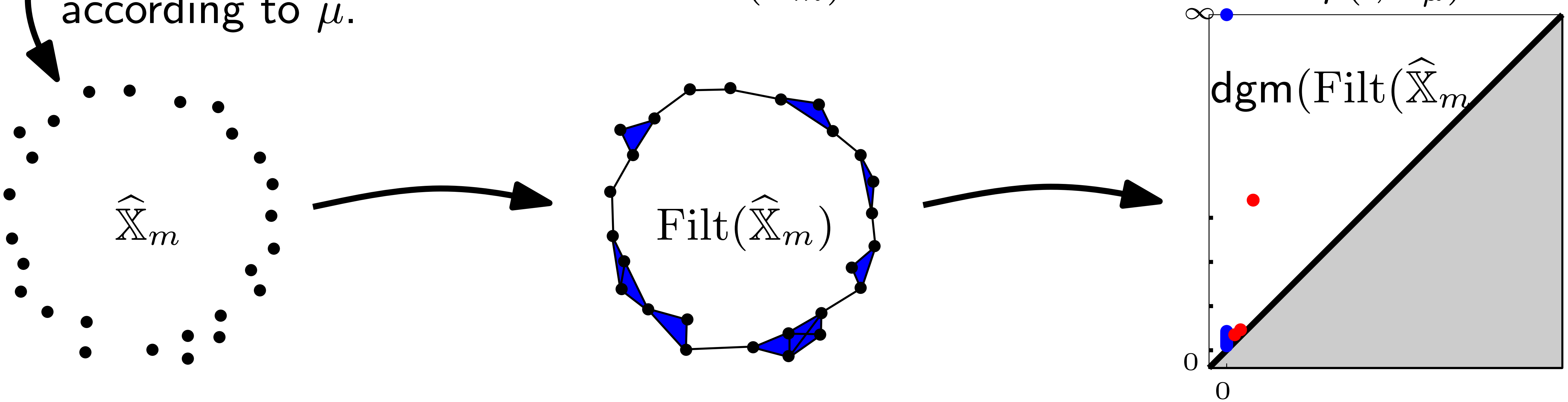


$(\mathbb{M}, \rho)$ metric space

μ a probability measure with compact support $\mathbb{X}_\mu$.

Examples:
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) = \mathrm{Rips}_\alpha(\widehat{\mathbb{X}}_m)$
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) = \check{\mathrm{C}}\mathrm{ech}_\alpha(\widehat{\mathbb{X}}_m)$
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) =$ sublevelset filtration of $\rho(., \mathbb{X}_\mu)$.

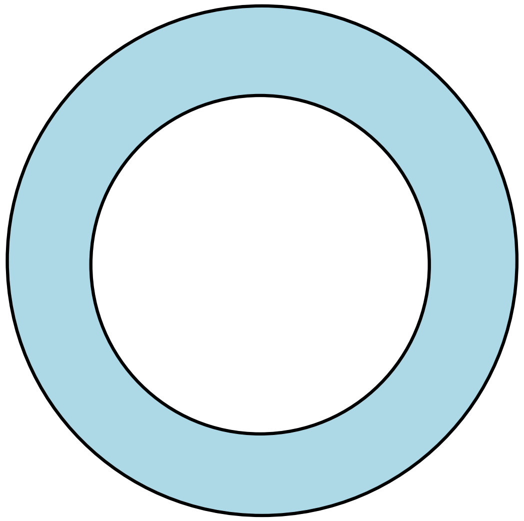
Sample m points according to μ.



Questions:

- Statistical properties of $\mathrm{dgm}(\mathrm{Filt}(\widehat{\mathbb{X}}_m))$? $\mathrm{dgm}(\mathrm{Filt}(\widehat{\mathbb{X}}_m)) \to ?$ as $m \to +\infty$?

Statistical setting

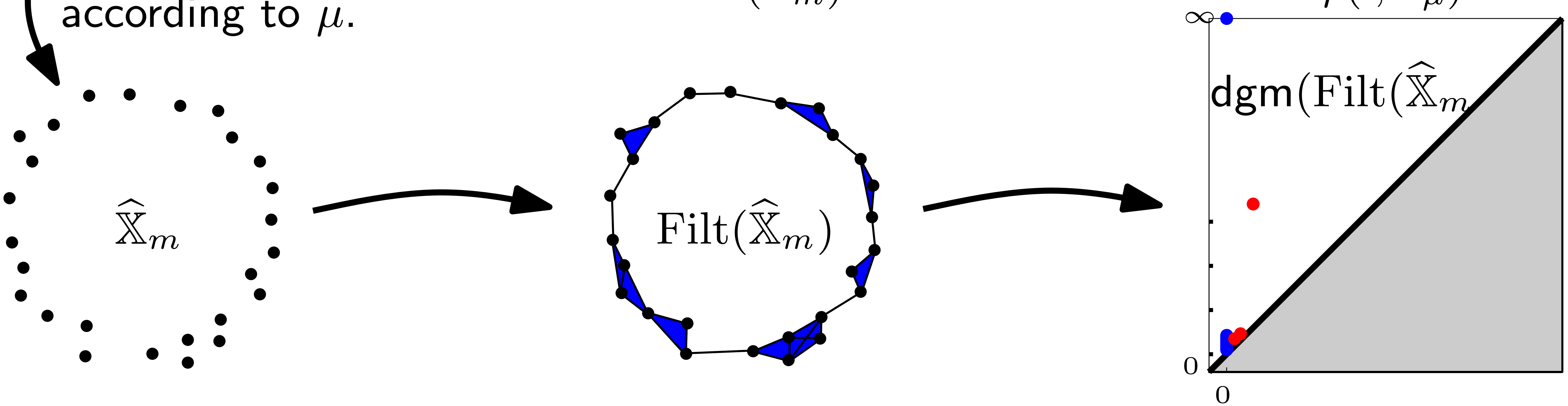


$(\mathbb{M}, \rho)$ metric space

μ a probability measure with compact support $\mathbb{X}_\mu$.

Examples:
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) = \mathrm{Rips}_\alpha(\widehat{\mathbb{X}}_m)$
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) = \check{\mathrm{C}}\mathrm{ech}_\alpha(\widehat{\mathbb{X}}_m)$
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) =$ sublevelset filtration of $\rho(., \mathbb{X}_\mu)$.

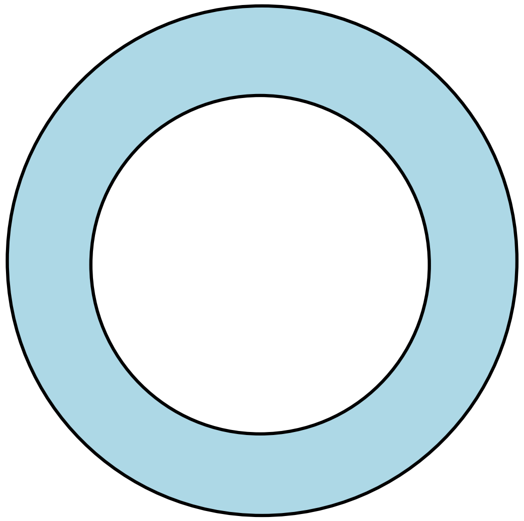
Sample m points according to μ.



Questions:

- Statistical properties of dgm$(\mathrm{Filt}(\widehat{\mathbb{X}}_m))$? dgm$(\mathrm{Filt}(\widehat{\mathbb{X}}_m)) \to ?$ as $m \to +\infty$?
- Can we do more statistics with persistence diagrams? What can be said about distributions of diagrams?

Statistical setting

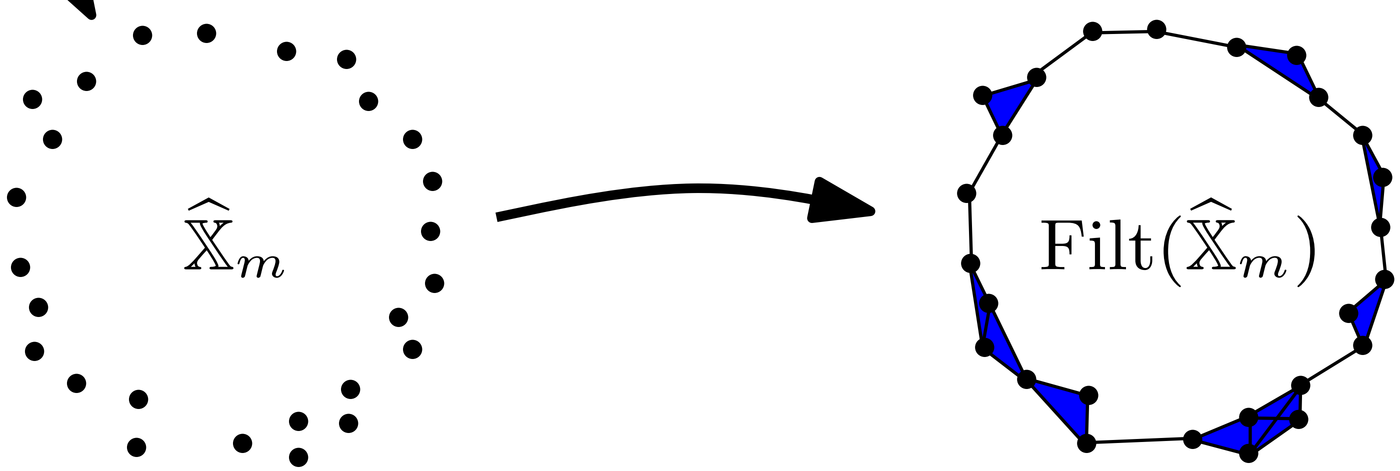


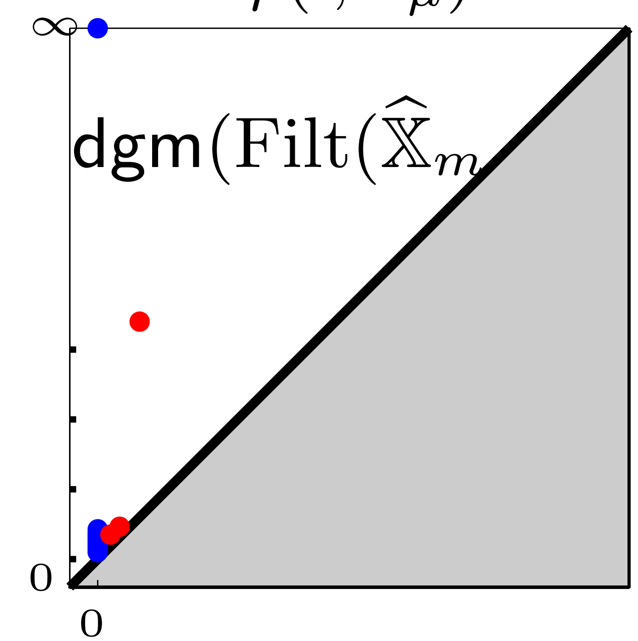
$(\mathbb{M}, \rho)$ metric space

μ a probability measure with compact support $\mathbb{X}_\mu$.

Examples:
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) = \mathrm{Rips}_\alpha(\widehat{\mathbb{X}}_m)$
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) = \check{\mathrm{C}}\mathrm{ech}_\alpha(\widehat{\mathbb{X}}_m)$
- $\mathrm{Filt}(\widehat{\mathbb{X}}_m) =$ sublevelset filtration of $\rho(., \mathbb{X}_\mu)$.

Sample m points according to μ.





Stability thm: $\mathrm{d_b}(\mathrm{dgm}(\mathrm{Filt}(\mathbb{X}_\mu)), \mathrm{dgm}(\mathrm{Filt}(\widehat{\mathbb{X}}_m))) \leq 2 d_{GH}(\mathbb{X}_\mu, \widehat{\mathbb{X}}_m)$

So, for any $\varepsilon > 0$,

$$\mathbb{P}\left(\mathrm{d_b}\left(\mathrm{dgm}(\mathrm{Filt}(\mathbb{X}_\mu)), \mathrm{dgm}(\mathrm{Filt}(\widehat{\mathbb{X}}_m))\right) > \varepsilon\right) \leq \mathbb{P}\left(d_{GH}(\mathbb{X}_\mu, \widehat{\mathbb{X}}_m) > \frac{\varepsilon}{2}\right)$$

Deviation inequality and rate of convergence

[C., Glisse, Labruère, Michel ICML'14 - JMLR'15]

For $a, b > 0$, μ satisfies the (a, b)-standard assumption if for any $x \in \mathbb{X}_\mu$ and any $r > 0$, we have $\mu(B(x, r)) \geq \min(ar^b, 1)$.

Deviation inequality and rate of convergence

[C., Glisse, Labruère, Michel ICML'14 - JMLR'15]

For $a, b > 0$, μ satisfies the (a,b)-standard assumption if for any $x \in \mathbb{X}_\mu$ and any $r > 0$, we have $\mu(B(x,r)) \geq \min(ar^b, 1)$.

Theorem: If μ satisfies the (a,b)-standard assumption, then for any $\varepsilon > 0$:

$$\mathbb{P}\left(\mathrm{d_b}\left(\mathsf{dgm}(\mathrm{Filt}(\mathbb{X}_\mu)), \mathsf{dgm}(\mathrm{Filt}(\widehat{\mathbb{X}}_m))\right) > \varepsilon\right) \leq \min(\frac{8^b}{a\varepsilon^b}\exp(-ma\varepsilon^b), 1).$$

Deviation inequality and rate of convergence

[C., Glisse, Labruère, Michel ICML'14 - JMLR'15]

For $a, b > 0$, μ satisfies the (a, b)-standard assumption if for any $x \in \mathbb{X}_\mu$ and any $r > 0$, we have $\mu(B(x, r)) \geq \min(ar^b, 1)$.

Theorem: If μ satisfies the (a, b)-standard assumption, then for any $\varepsilon > 0$:

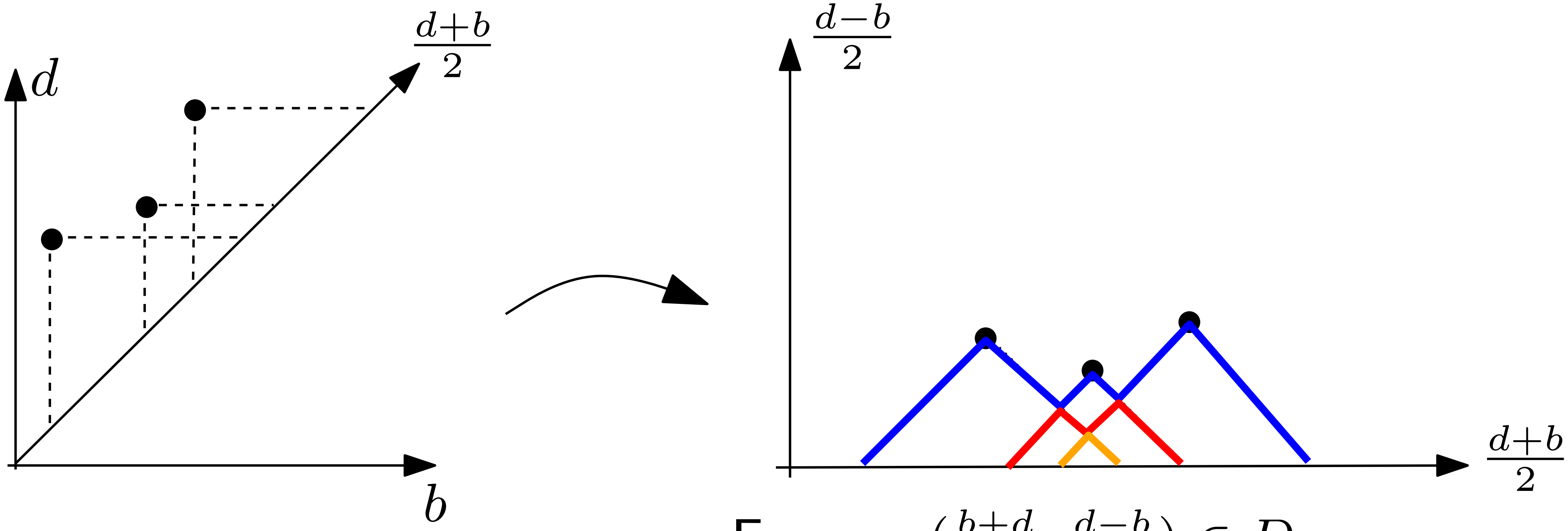
$$\mathbb{P}\left(\mathrm{d_b}\left(\mathsf{dgm}(\mathrm{Filt}(\mathbb{X}_\mu)), \mathsf{dgm}(\mathrm{Filt}(\widehat{\mathbb{X}}_m))\right) > \varepsilon\right) \leq \min(\frac{8^b}{a\varepsilon^b}\exp(-ma\varepsilon^b), 1).$$

Corollary: Let $\mathcal{P}(a, b, \mathbb{M})$ be the set of (a, b)-standard proba measures on $\mathbb{M}$. Then:

$$\sup_{\mu \in \mathcal{P}(a,b,\mathbb{M})} \mathbb{E}\left[\mathrm{d_b}(\mathsf{dgm}(\mathrm{Filt}(\mathbb{X}_\mu)), \mathsf{dgm}(\mathrm{Filt}(\widehat{\mathbb{X}}_m)))\right] \leq C\left(\frac{\ln m}{m}\right)^{1/b}$$

where the constant C only depends on a and b (not on $\mathbb{M}$!). Moreover, the upper bound is tight (in a minimax sense)!

Persistence landscapes



$D = \{(\frac{d_i+b_i}{2}, \frac{d_i+b_i}{2})\} i \in I$

For $p = (\frac{b+d}{2}, \frac{d-b}{2}) \in D$,

$$\Lambda_p(t) = \begin{cases} t - b & t \in [b, \frac{b+d}{2}] \\ d - t & t \in (\frac{b+d}{2}, d] \\ 0 & \text{otherwise.} \end{cases}$$

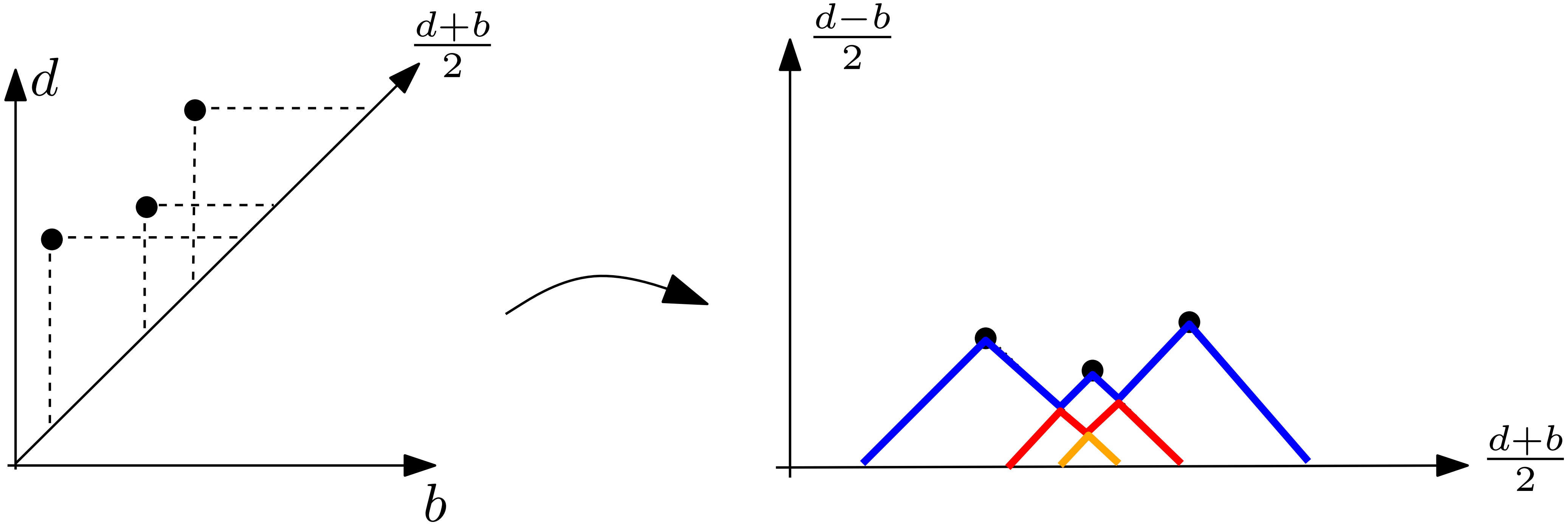
Persistence landscape [Bubenik 2012]:

$$\lambda_D(k,t) = \underset{p \in \mathrm{dgm}}{\mathrm{kmax}}\, \Lambda_p(t), \quad t \in \mathbb{R}, k \in \mathbb{N},$$

where kmax is the kth largest value in the set.

Many other ways to “linearize” persistence diagrams: intensity functions, image persistence, kernels,...

Persistence landscapes



Persistence landscape [Bubenik 2012]:

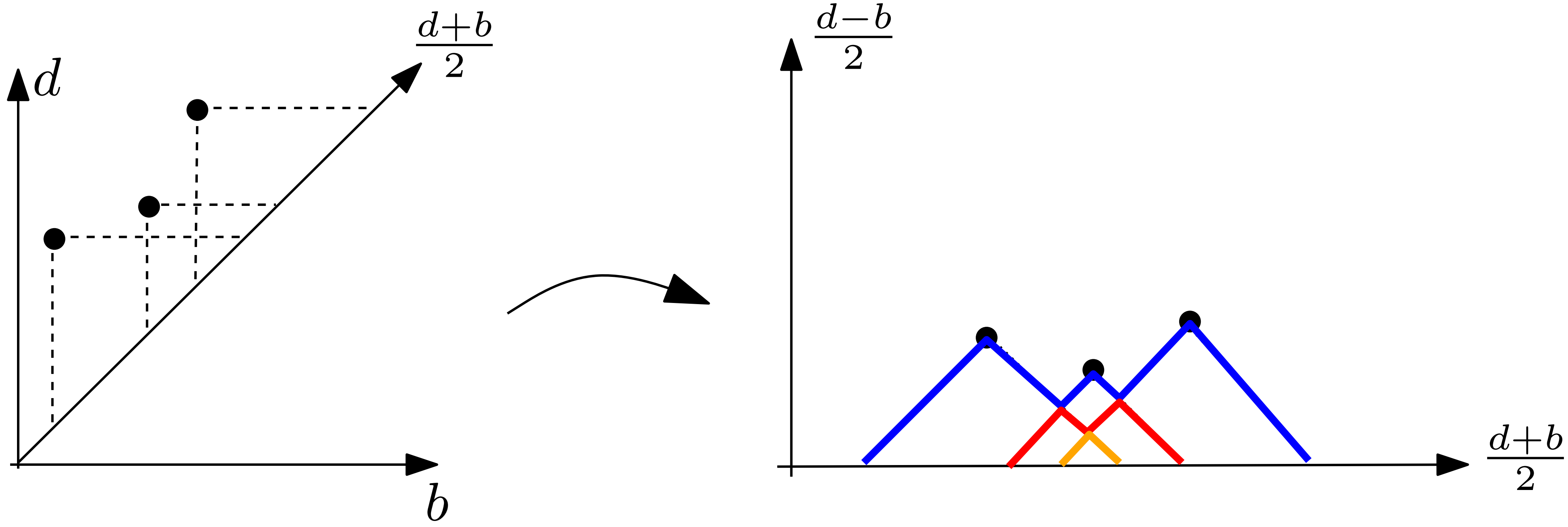
$$\lambda_D(k,t) = \underset{p \in \mathsf{dgm}}{\mathrm{kmax}}\ \Lambda_p(t), \quad t \in \mathbb{R}, k \in \mathbb{N},$$

Properties

- For any $t \in \mathbb{R}$ and any $k \in \mathbb{N}$, $0 \leq \lambda_D(k,t) \leq \lambda_D(k+1,t)$.
- For any $t \in \mathbb{R}$ and any $k \in \mathbb{N}$, $|\lambda_D(k,t) - \lambda_{D'}(k,t)| \leq d_B(D,D')$ where $d_B(D,D')$ denotes the bottleneck distance between D and D'.

stability properties of persistence landscapes

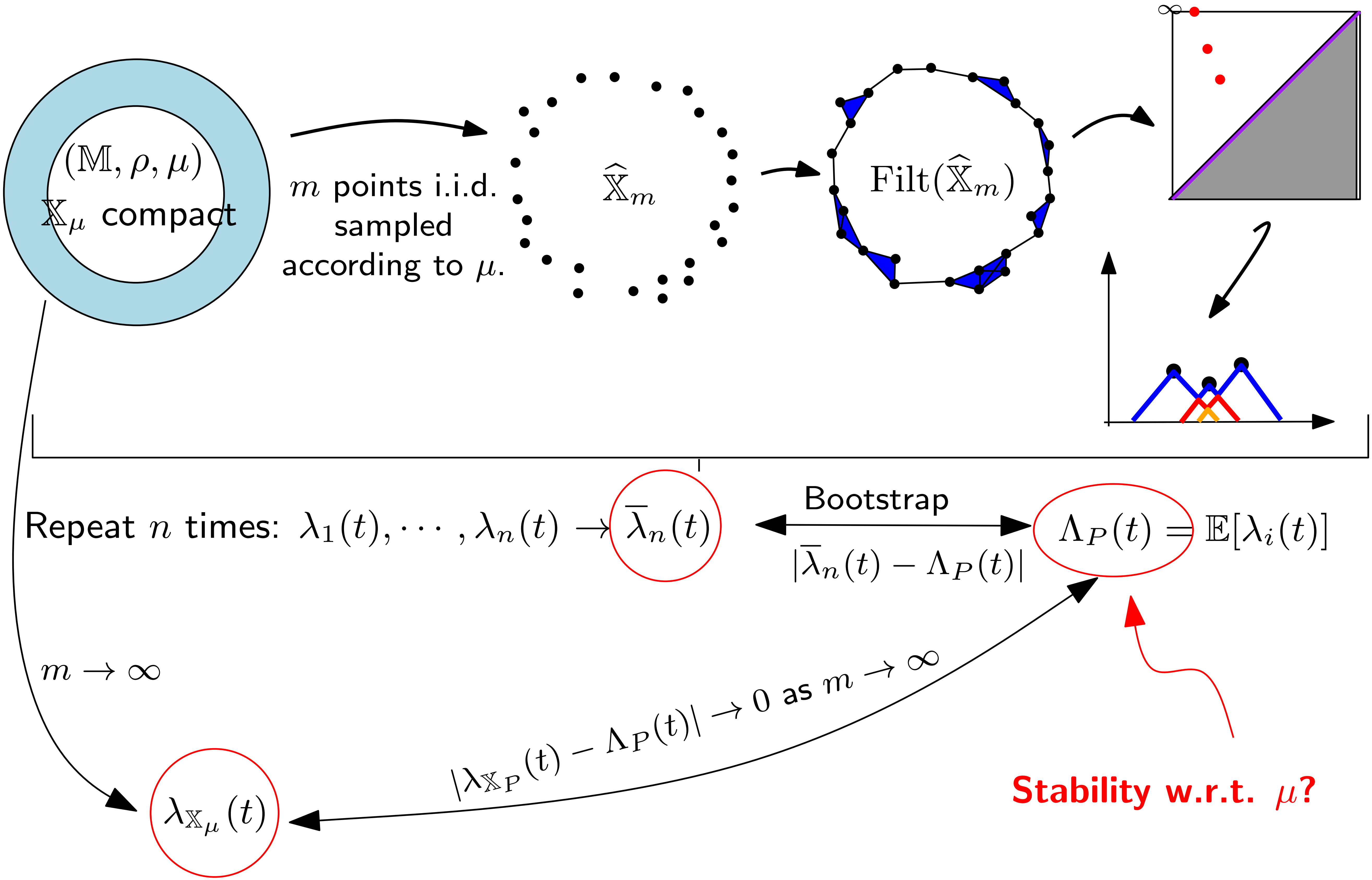
Persistence landscapes



- Persistence encoded as an element of a functional space (vector space!).
- Expectation of distribution of landscapes is well-defined and can be approximated from average of sampled landscapes.
- process point of view: convergence results and convergence rates $\rightarrow$ confidence intervals can be computed using bootstrap.

[C., Fasy, Lecci, Rinaldo, Wasserman SoCG 2014]

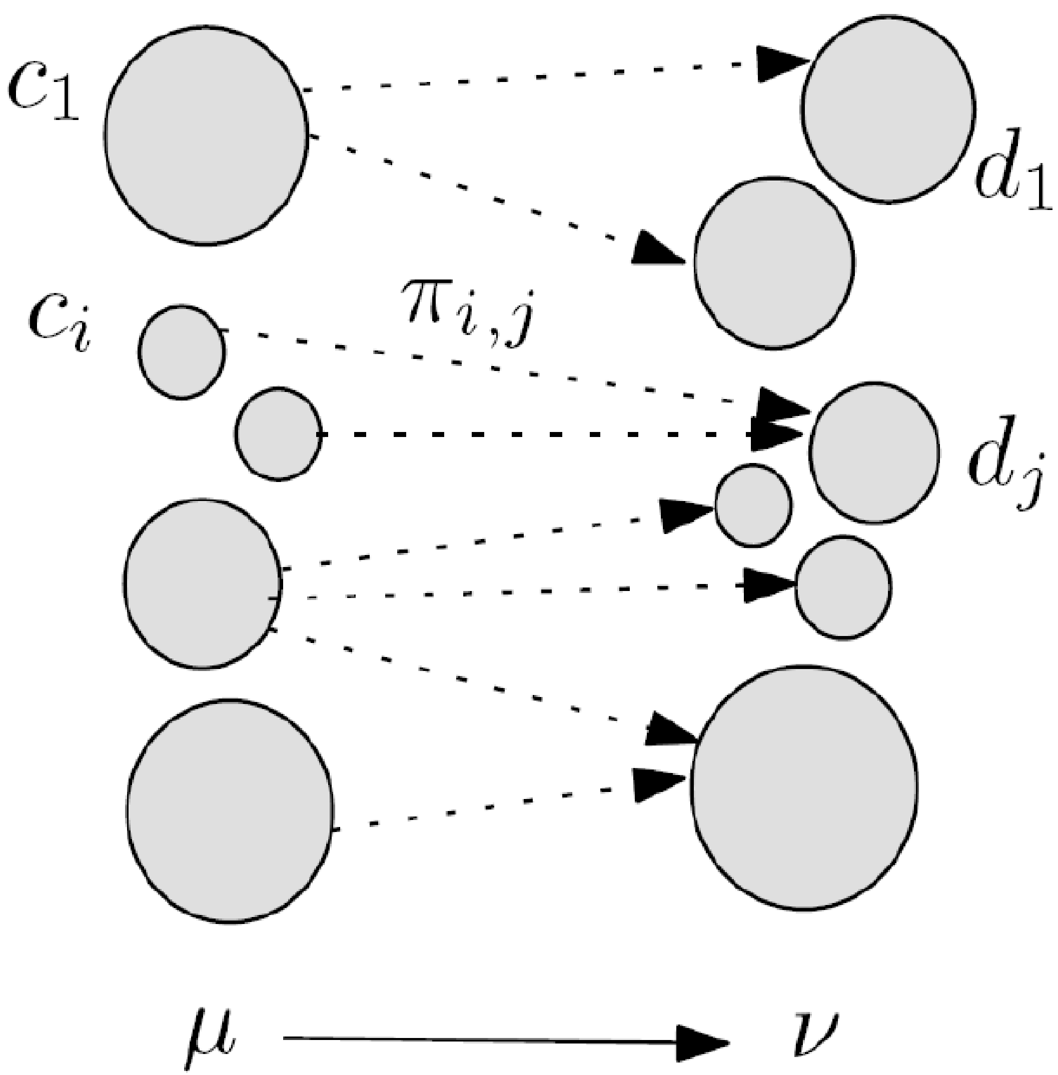
To summarize



Wasserstein distance

Let $(\mathbb{M}, \rho)$ be a metric space and let μ, ν be probability measures on $\mathbb{M}$ with finite p-moments ($p \geq 1$).

"The" Wasserstein distance $W_p(\mu, \nu)$ quantifies the optimal cost of pushing μ onto ν, the cost of moving a small mass dx from x to y being $\rho(x, y)^p dx$.



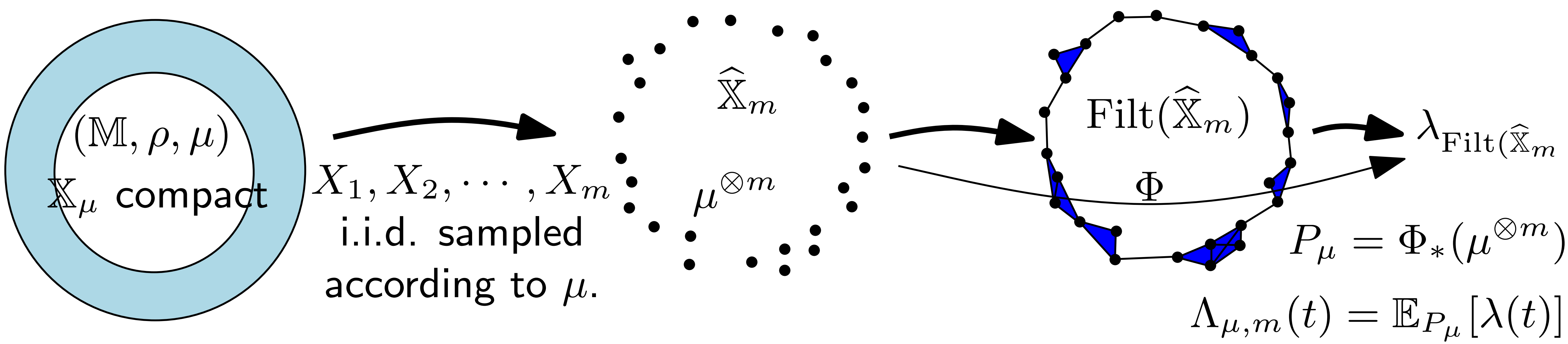
- Transport plan: Π a proba measure on $M \times M$ such that $\Pi(A \times \mathbb{R}^d) = \mu(A)$ and $\Pi(\mathbb{R}^d \times B) = \nu(B)$ for any borelian sets $A, B \subset M$.
- Cost of a transport plan:

$$C(\Pi) = \left(\int_{M \times M} \rho(x, y)^p d\Pi(x, y) \right)^{\frac{1}{p}}$$

- $W_p(\mu, \nu) = \inf_\Pi C(\Pi)$

(Sub)sampling and stability of expected landscapes

[C., Fasy, Lecci, Michel, Rinaldo, Wasserman ICML 2015]



Theorem: Let $(\mathbb{M}, \rho)$ be a metric space and let μ, ν be proba measures on $\mathbb{M}$ with compact supports. We have

$$\|\Lambda_{\mu,m} - \Lambda_{\nu,m}\|_\infty \leq m^{\frac{1}{p}} W_p(\mu, \nu)$$

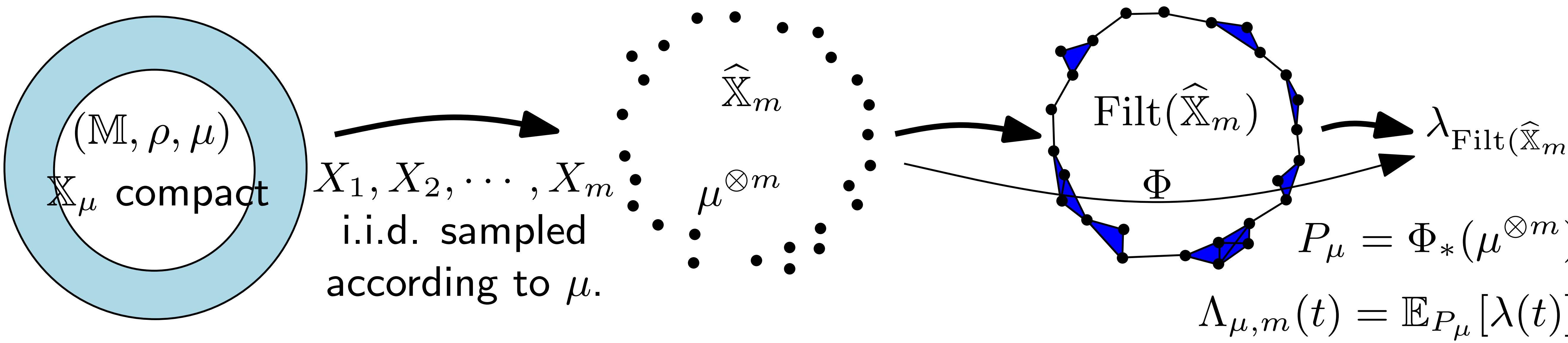
where W_p denotes the Wasserstein distance with cost function $\rho(x,y)^p$.

Remarks:

- similar results by Blumberg et al (2014) in the (Gromov-)Prokhorov metric (for distributions, not for expectations) ;
- Extended to point process setting y L. Decreusefond et al;
- $m^{\frac{1}{p}}$ cannot be replaced by a constant.

(Sub)sampling and stability of expected landscapes

[C., Fasy, Lecci, Michel, Rinaldo, Wasserman ICML 2015]



Theorem: Let $(\mathbb{M}, \rho)$ be a metric space and let μ, ν be proba measures on $\mathbb{M}$ with compact supports. We have

$$\|\Lambda_{\mu,m} - \Lambda_{\nu,m}\|_\infty \leq m^{\frac{1}{p}} W_p(\mu, \nu)$$

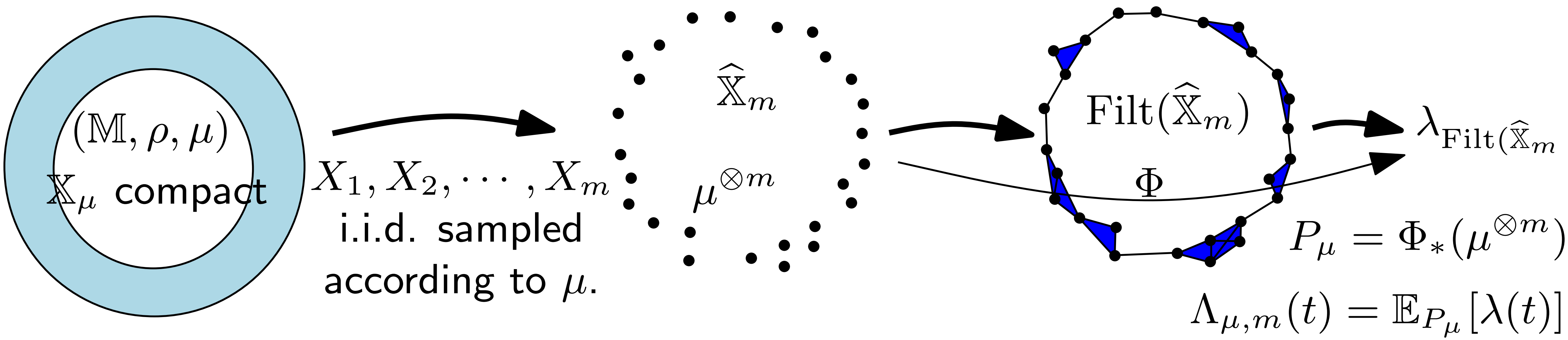
where W_p denotes the Wasserstein distance with cost function $\rho(x, y)^p$.

Consequences:

- Subsampling: efficient and easy to parallelize algorithm to infer topol. information from huge data sets.
- Robustness to outliers.
- R package TDA +Gudhi library: https://project.inria.fr/gudhi/software/

(Sub)sampling and stability of expected landscapes

[C., Fasy, Lecci, Michel, Rinaldo, Wasserman ICML 2015]



Theorem: Let $(\mathbb{M}, \rho)$ be a metric space and let μ, ν be proba measures on $\mathbb{M}$ with compact supports. We have

$$\|\Lambda_{\mu,m} - \Lambda_{\nu,m}\|_\infty \leq m^{\frac{1}{p}} W_p(\mu, \nu)$$

where W_p denotes the Wasserstein distance with cost function $\rho(x, y)^p$.

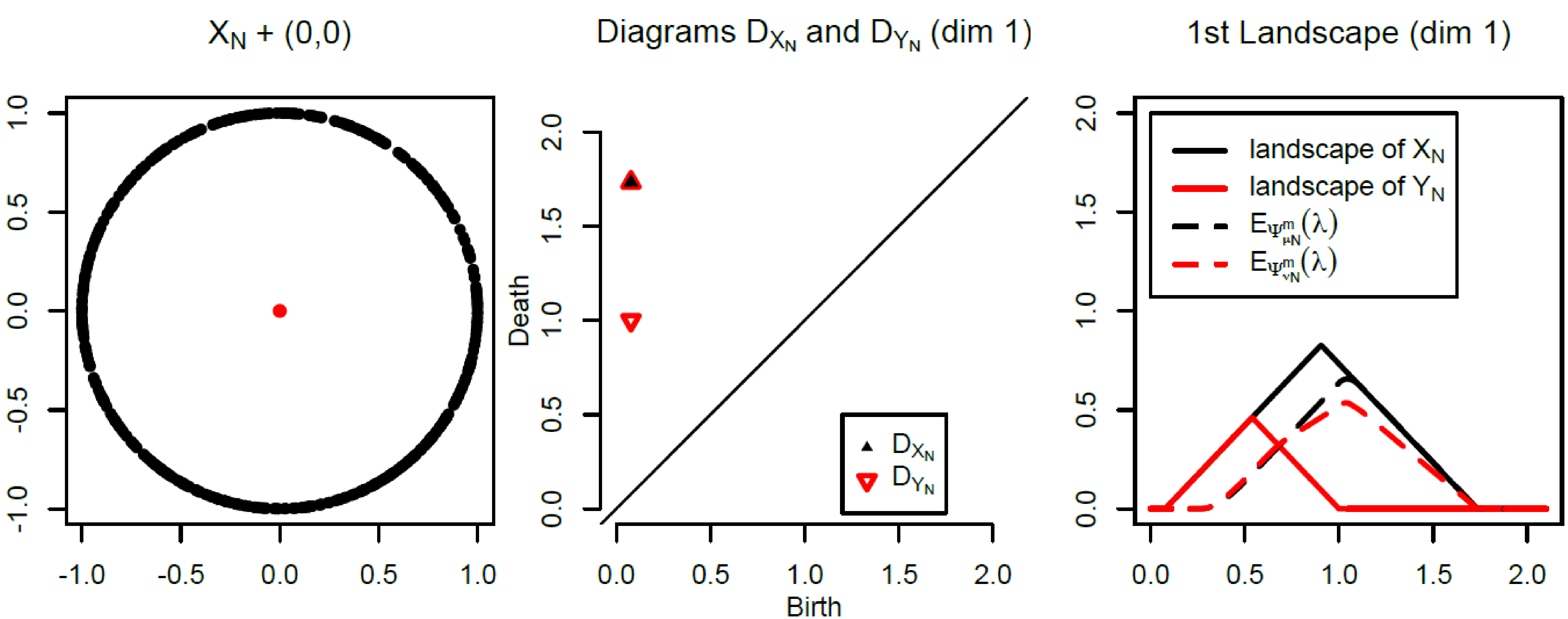
Proof:

1. $W_p(\mu^{\otimes m}, \nu^{\otimes m}) \leq m^{\frac{1}{p}} W_p(\mu, \nu)$
2. $W_p(P_\mu, P_\nu) \leq W_p(\mu^{\otimes m}, \nu^{\otimes m})$ (stability of persistence!)
3. $\|\Lambda_{\mu,m} - \Lambda_{\nu,m}\|_\infty \leq W_p(P_\mu, P_\nu)$ (Jensen's inequality)

(Sub)sampling and stability of expected landscapes

[C., Fasy, Lecci, Michel, Rinaldo, Wasserman ICML 2015]

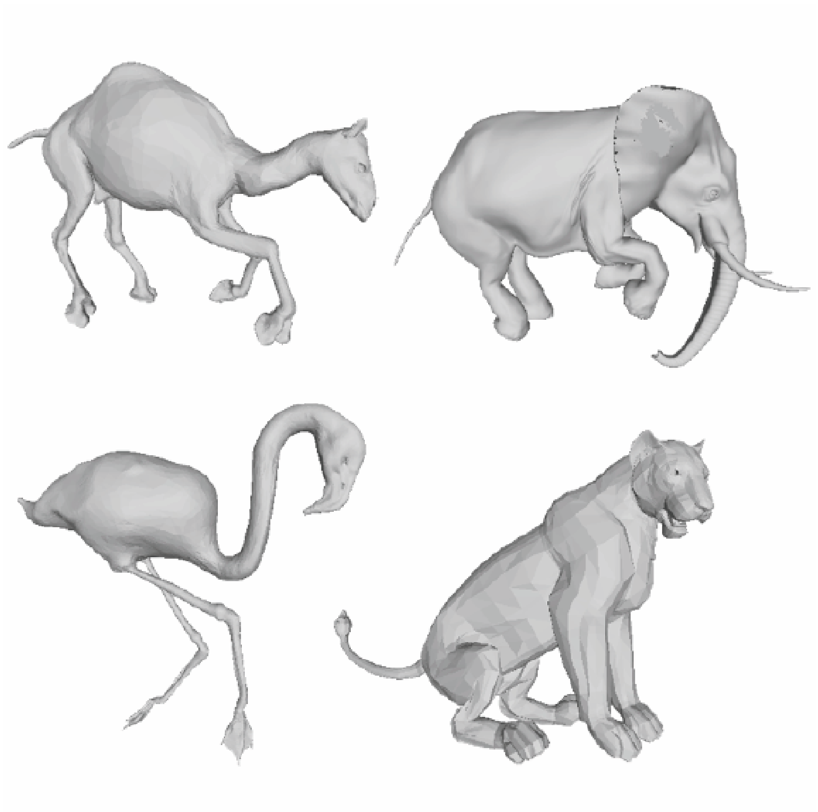
Example: Circle with one outlier.

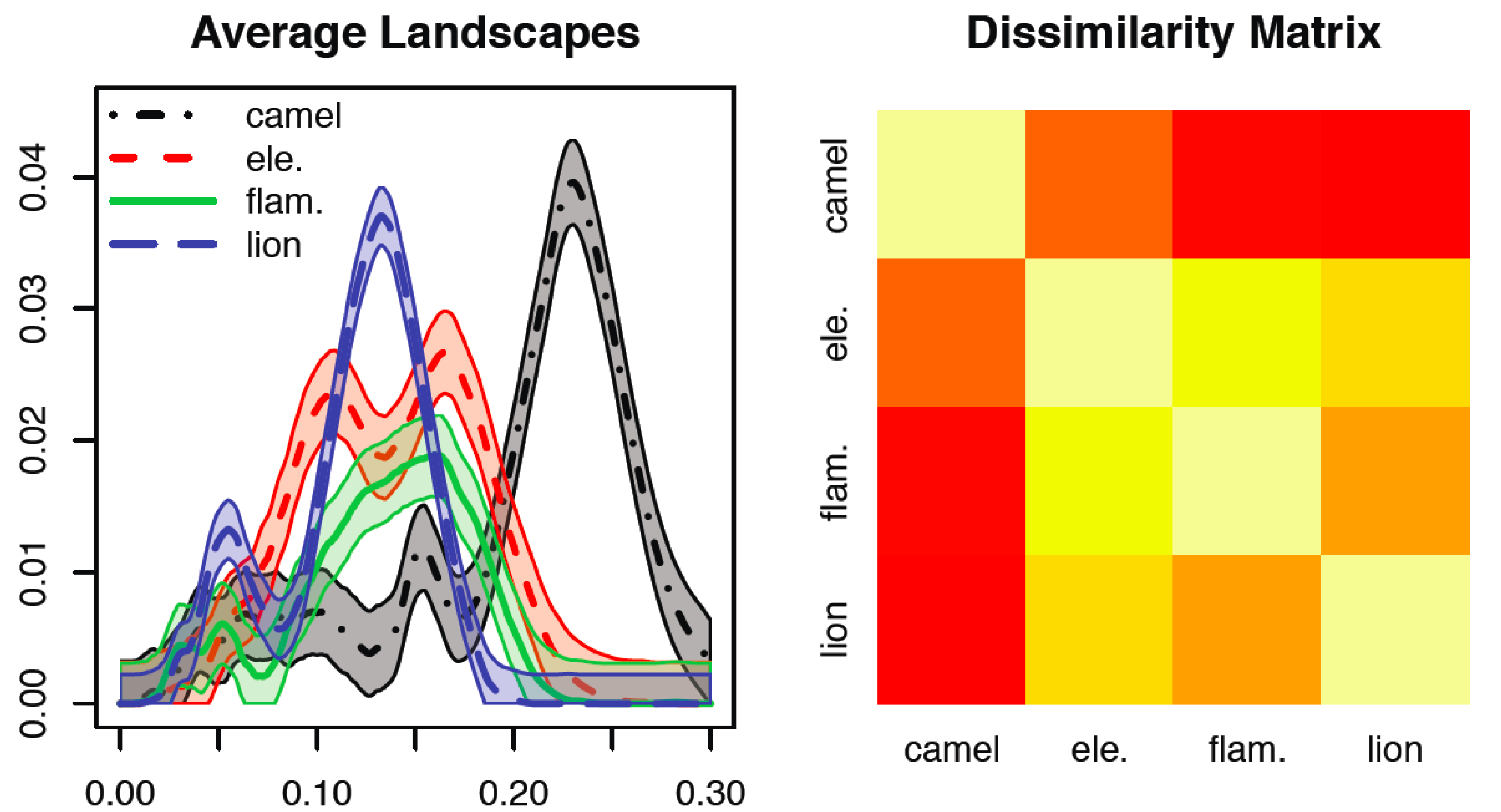


(Sub)sampling and stability of expected landscapes

[C., Fasy, Lecci, Michel, Rinaldo, Wasserman ICML 2015]

Example: 3D shapes



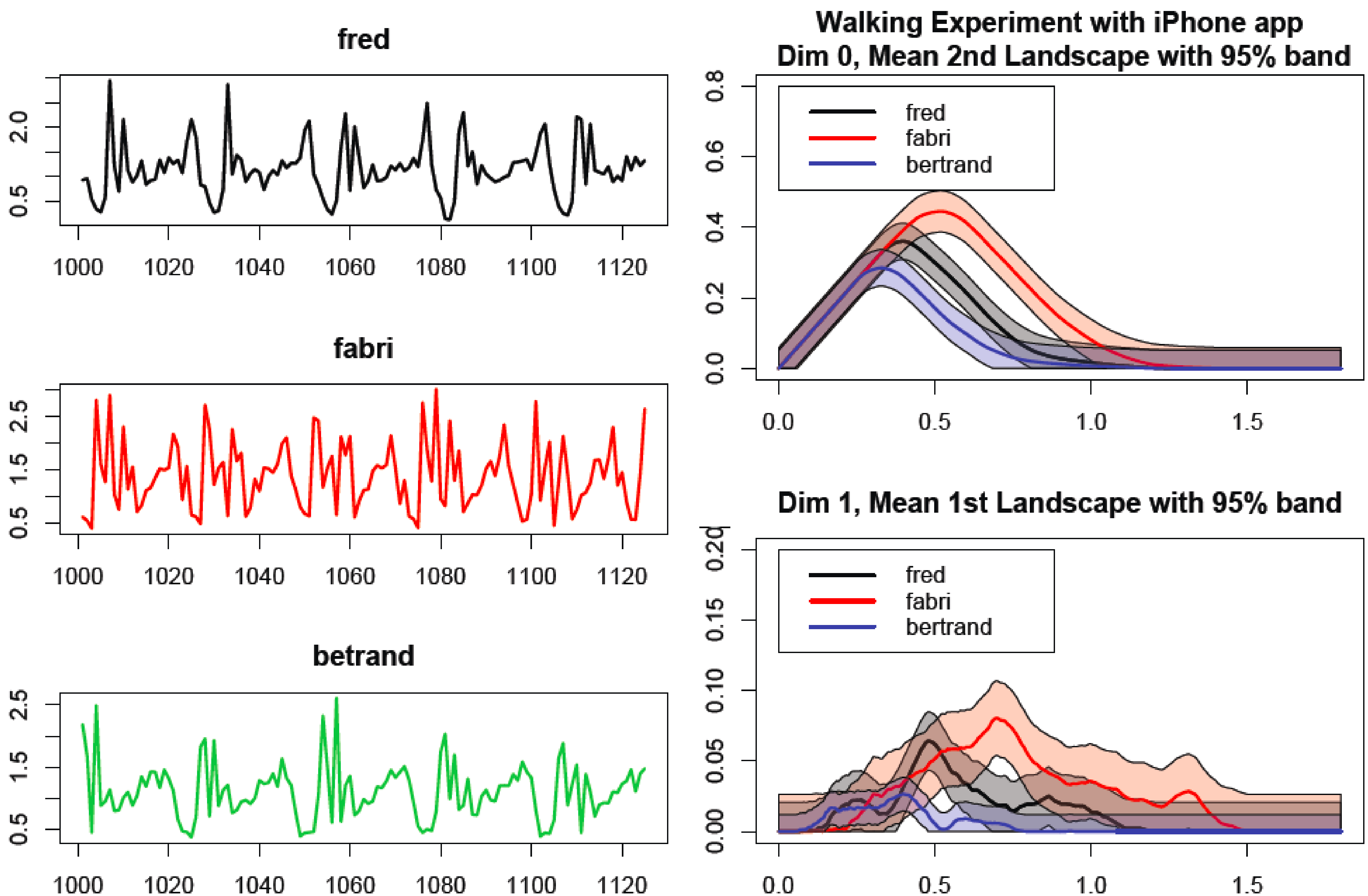


From $n = 100$ subsamples of size $m = 300$

(Sub)sampling and stability of expected landscapes

[C., Fasy, Lecci, Michel, Rinaldo, Wasserman ICML 2015]

(Toy) Example: Accelerometer data from smartphone.



- spatial time series (accelerometer data from the smarphone of users).
- no registration/calibration preprocessing step needed to compare!

Thank you for your attention!

Collaborators: T. Bonis, V. de Silva, B. Fasy, D. Cohen-Steiner, M. Glisse, L. Guibas, C. Labruère, F. Lecci, C. Li, F. Memoli, B. Michel, S. Oudot, M. Ovsjanikov, A. Rinaldo, P. Skraba, L. Wasserman

Software:

- The Gudhi library (C++/Python): https://project.inria.fr/gudhi/software/
- R package TDA