

Journées de Géométrie Algorithmique, January 2009

Analysis of Scalar Fields over Point Cloud Data

Frédéric Chazal, Leonidas J. Guibas, Steve Y. Oudot, Primož Skraba

Géometrica Group
INRIA Saclay

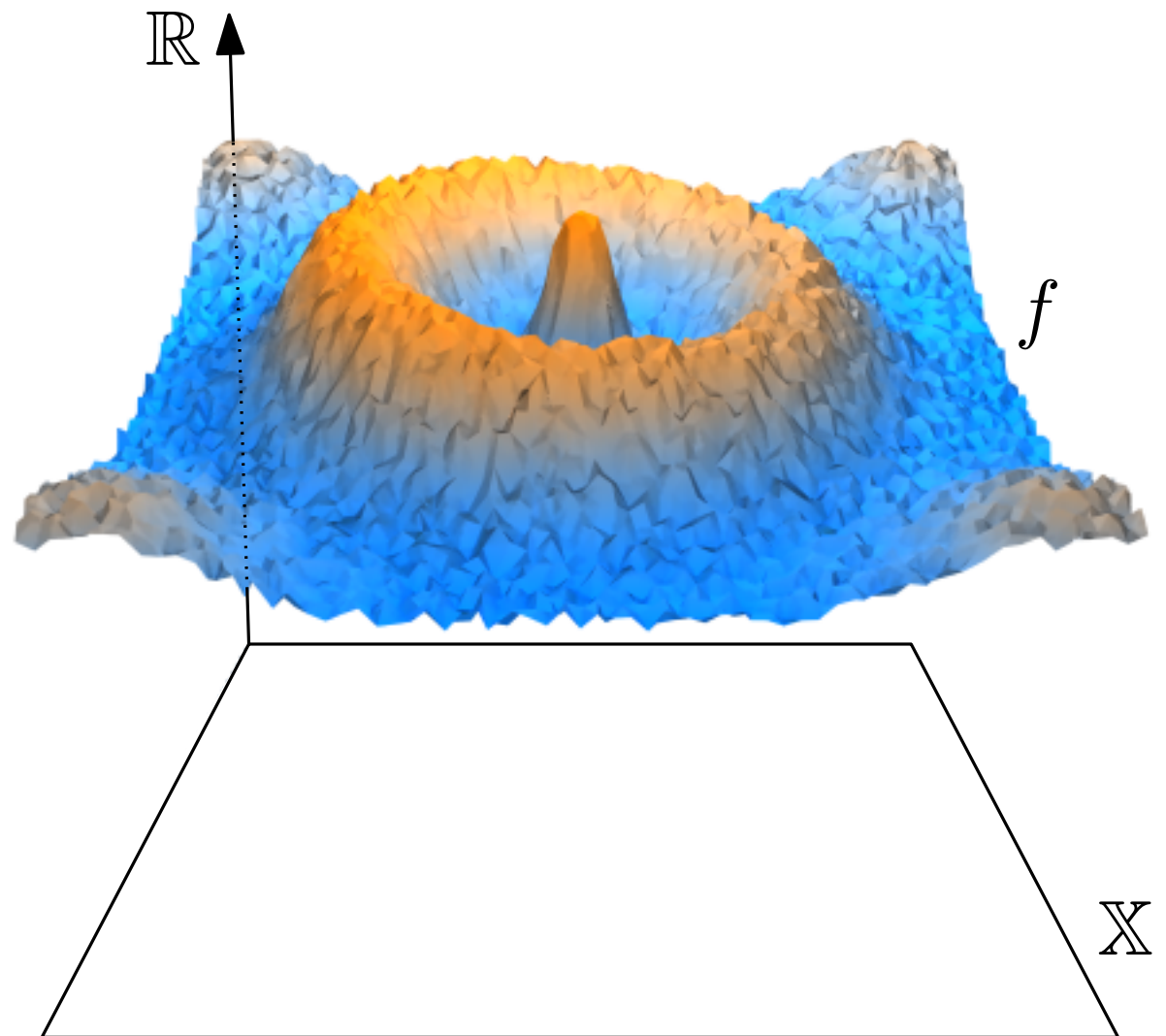


Computer Science Department
Stanford University



Scalar Field Analysis

Setting: $\mathbb{X}$ topological space, $f : \mathbb{X} \rightarrow \mathbb{R}$



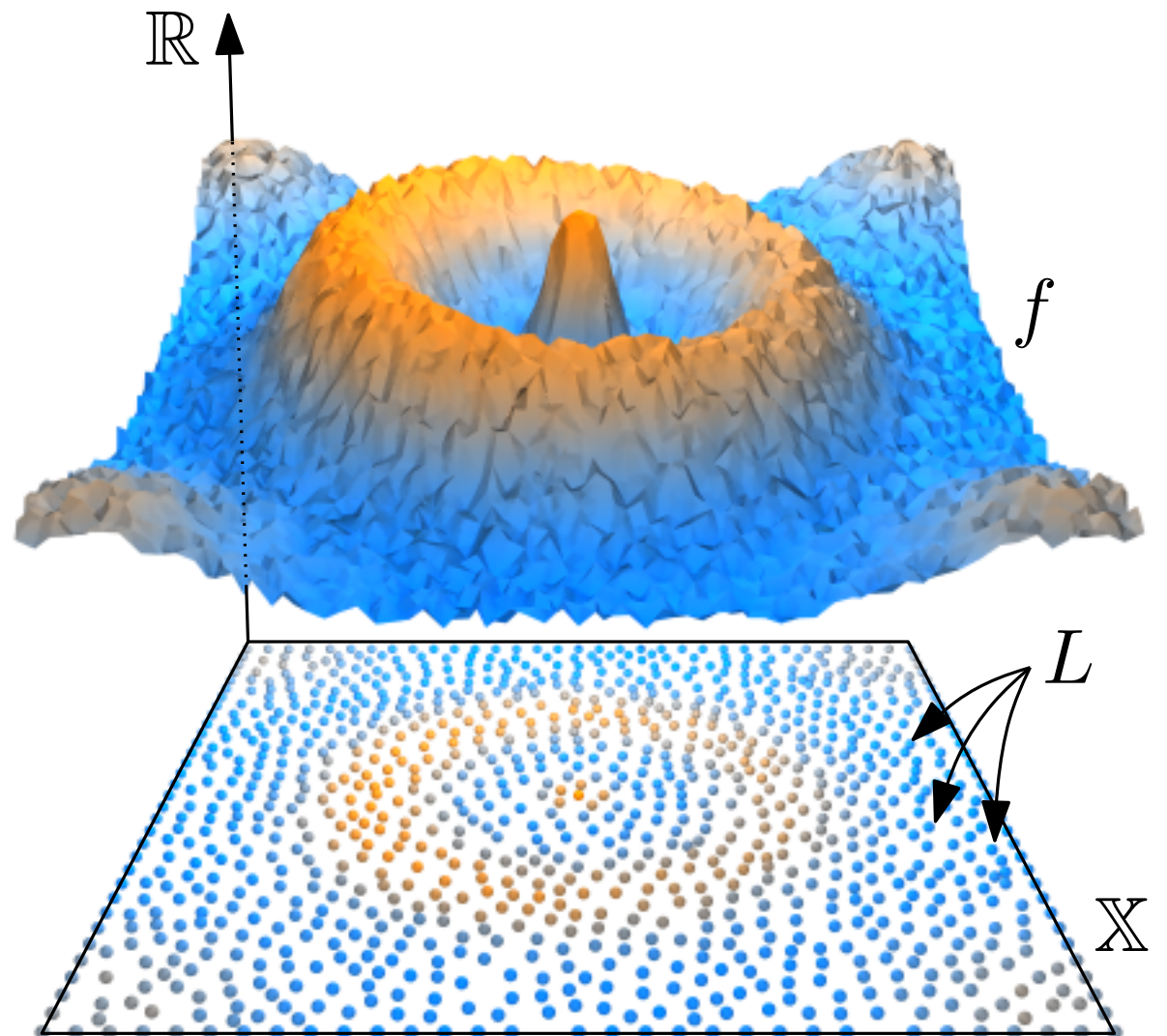
Scalar Field Analysis

Setting: $\mathbb{X}$ topological space, $f : \mathbb{X} \rightarrow \mathbb{R}$

Input: A finite sampling L of $\mathbb{X}$, the values of f at the sample points

Goal: Analyze landscape of $\text{graph}(f)$:

- *prominent peaks/valleys*



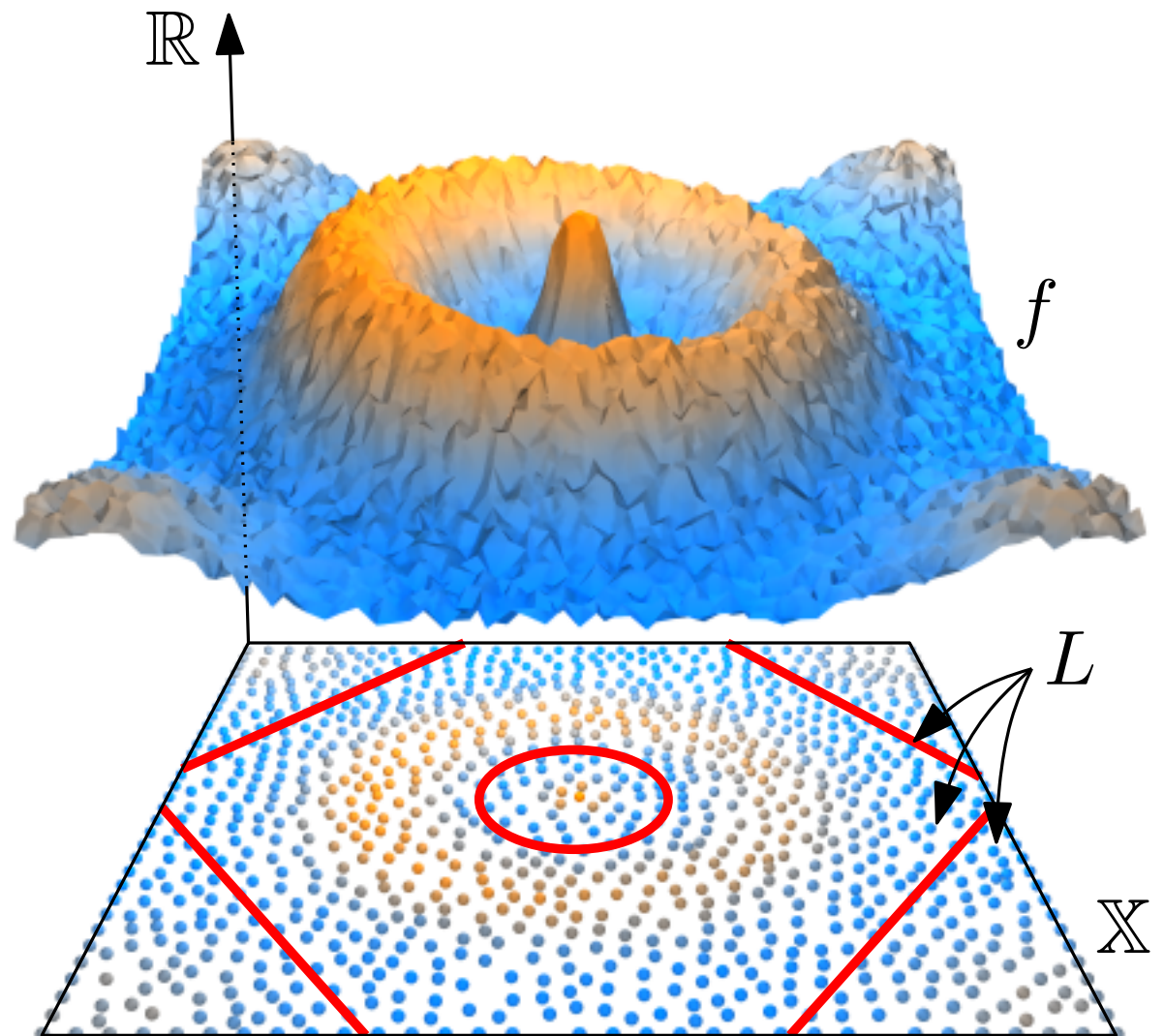
Scalar Field Analysis

Setting: $\mathbb{X}$ topological space, $f : \mathbb{X} \rightarrow \mathbb{R}$

Input: A finite sampling L of $\mathbb{X}$, the values of f at the sample points

Goal: Analyze landscape of $\text{graph}(f)$:

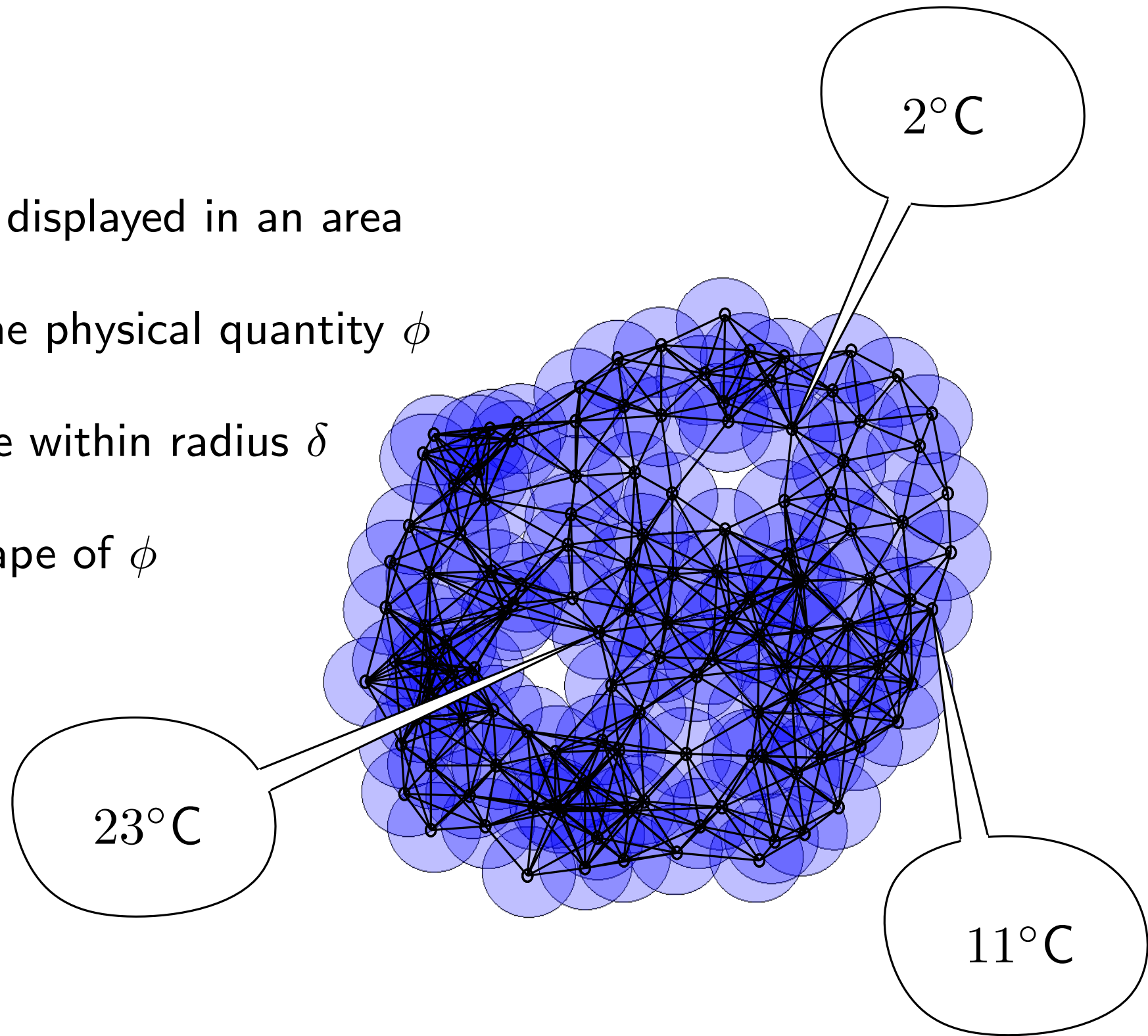
- *prominent* peaks/valleys
- basins of attraction



Motivating Applications

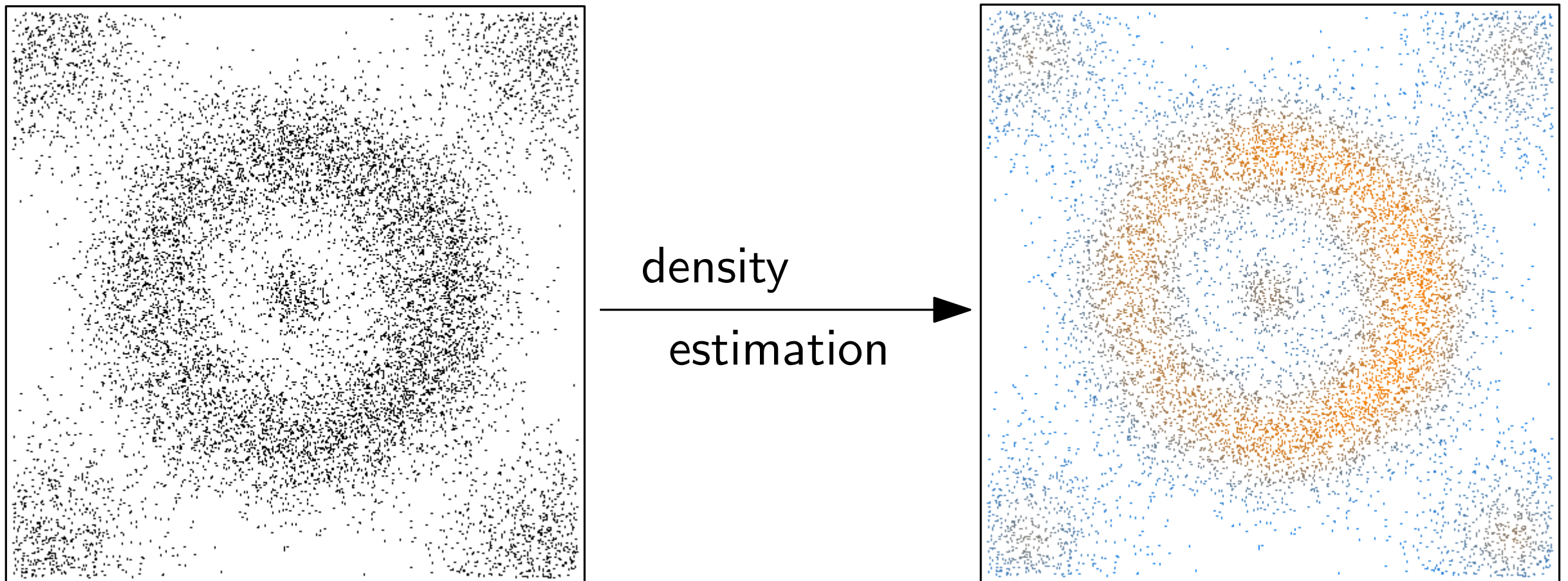
- Sensor networks:
 - collection of sensors displayed in an area
 - sensors measure same physical quantity ϕ
 - sensors communicate within radius δ

Goal: Analyze landscape of ϕ



Motivating Applications

- Sensor networks:
- Unsupervised learning:
 - data points drawn at random from some unknown density distribution f
 - approximate f through some density estimator $\hat{f}$
 - cluster data points according to *prominent* basins of attraction of $\hat{f}$

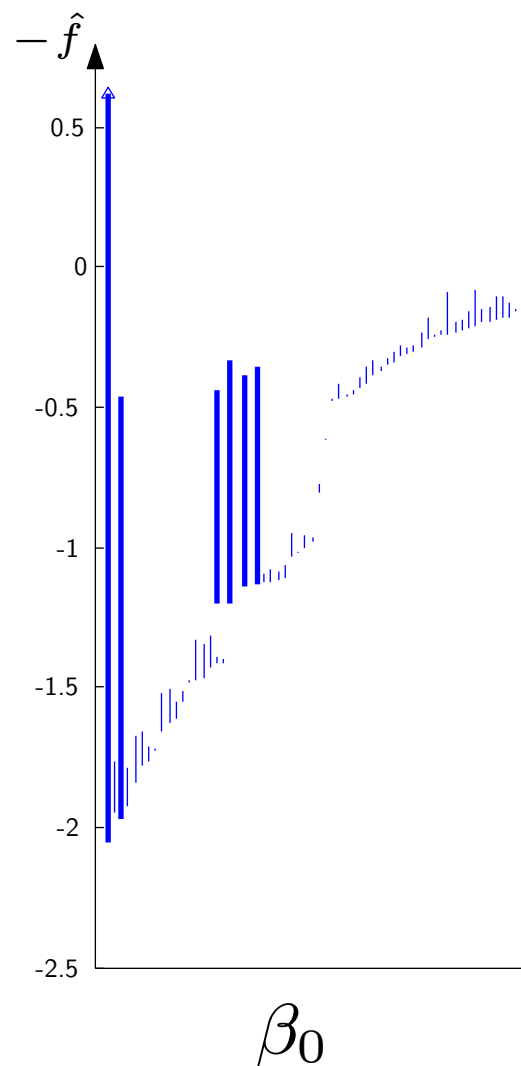


Persistence-Based Approach

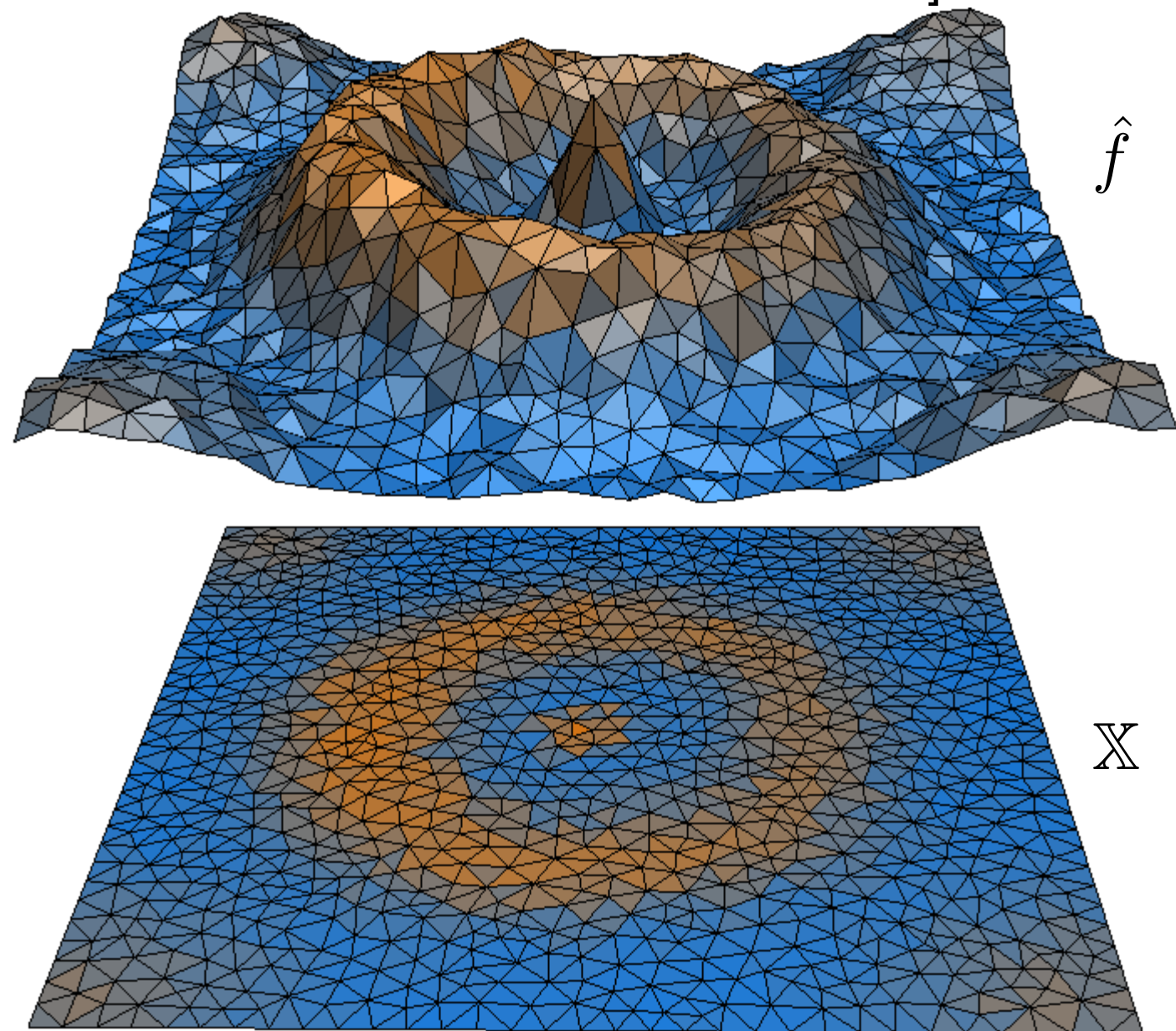
Assumptions: $\mathbb{X}$ **triangulated** space, $f : \mathbb{X} \rightarrow \mathbb{R}$ Lipschitz continuous

→ build PL approximation $\hat{f}$ of f

→ apply persistence algo. to $\pm \hat{f}$ [Edelsbrunner, Letscher, Zomorodian '00]



(6 prominent peaks)

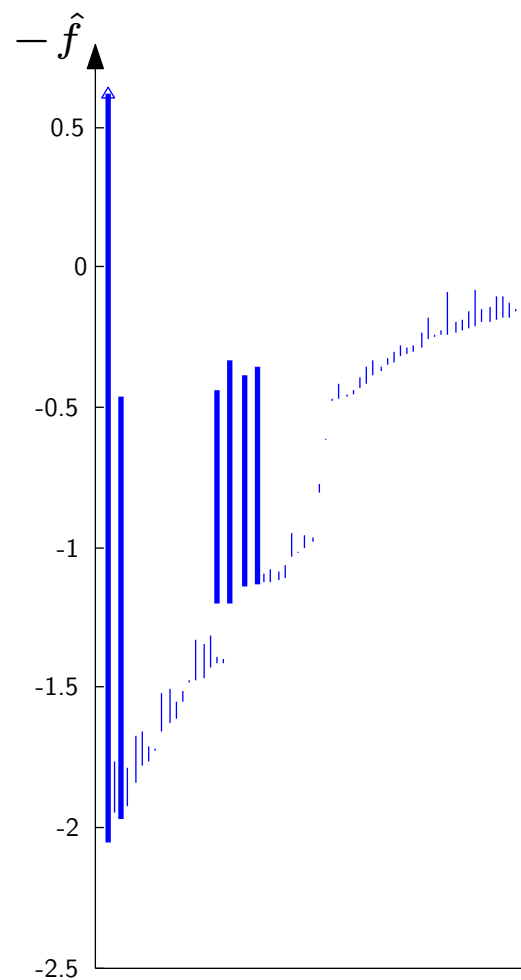


Persistence-Based Approach

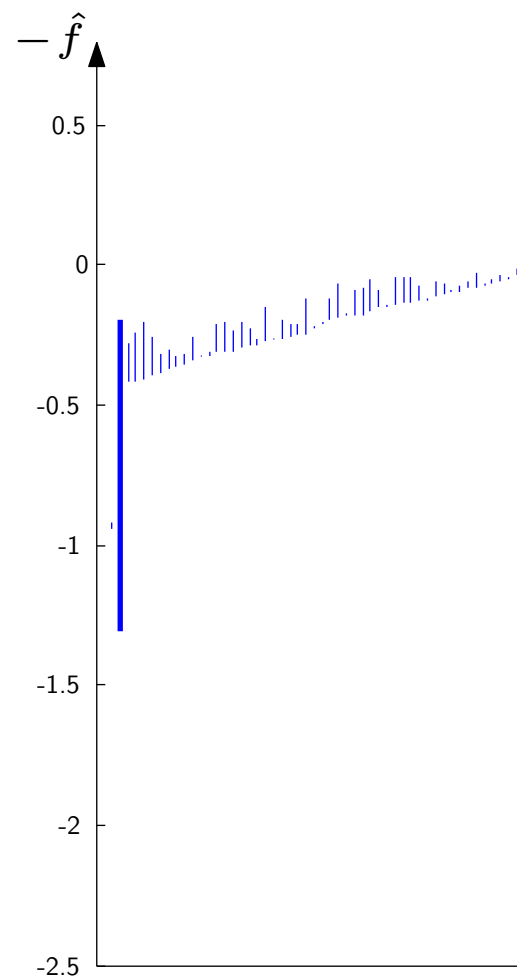
Assumptions: $\mathbb{X}$ **triangulated** space, $f : \mathbb{X} \rightarrow \mathbb{R}$ Lipschitz continuous

→ build PL approximation $\hat{f}$ of f

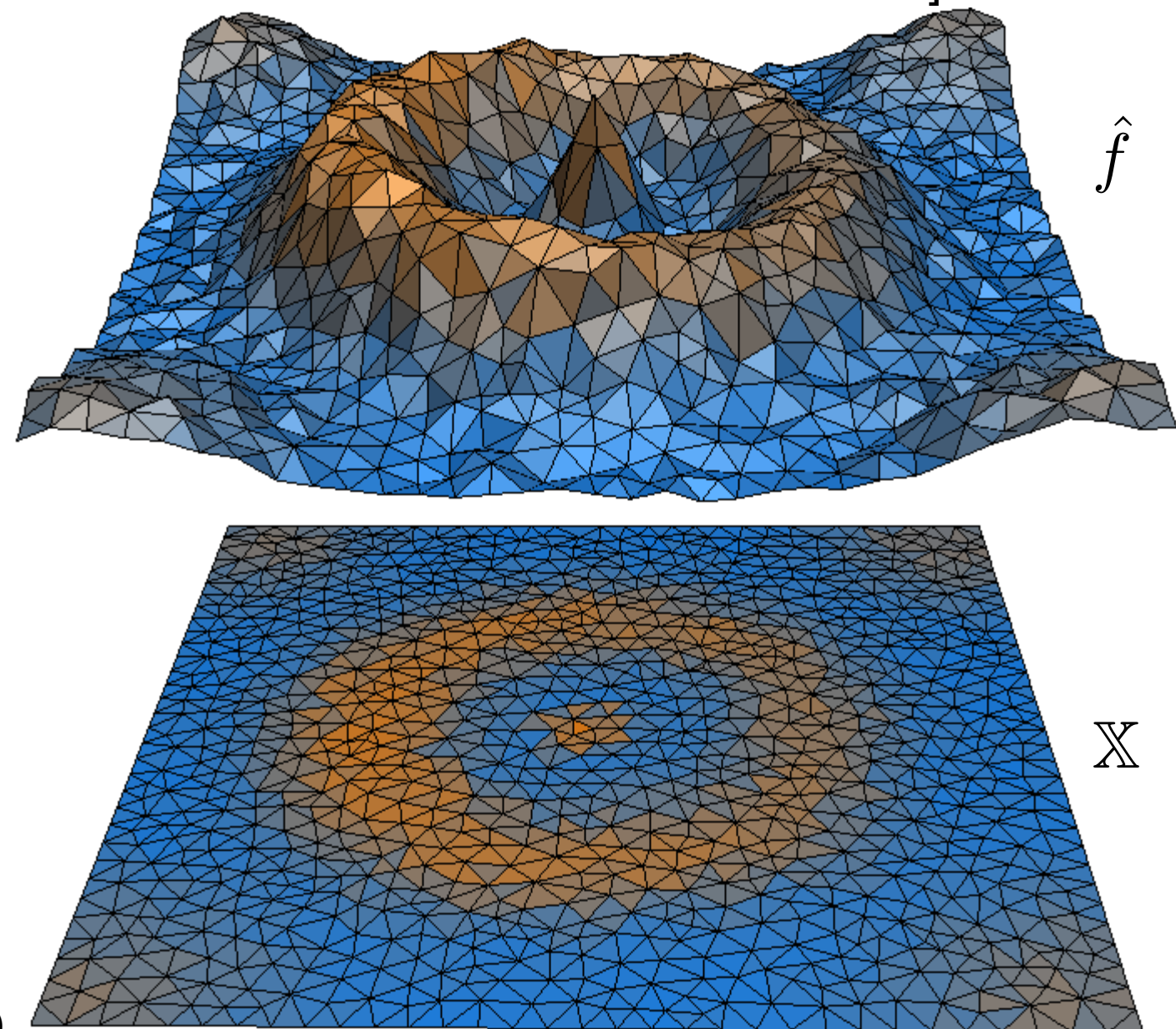
→ apply persistence algo. to $\pm \hat{f}$ [Edelsbrunner, Letscher, Zomorodian '00]



β_0
(6 prominent peaks)

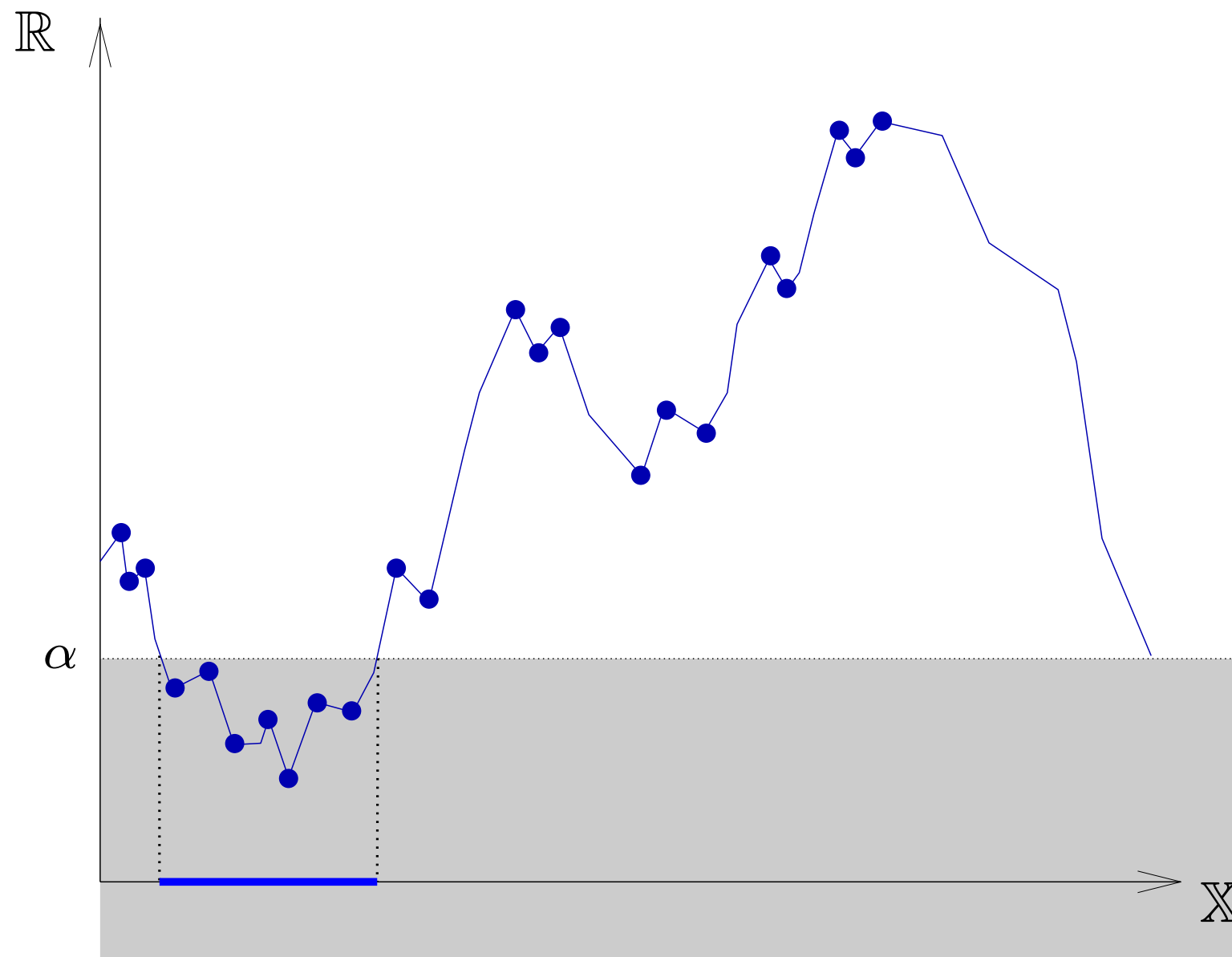


β_1
(ring-shaped basin of attraction)



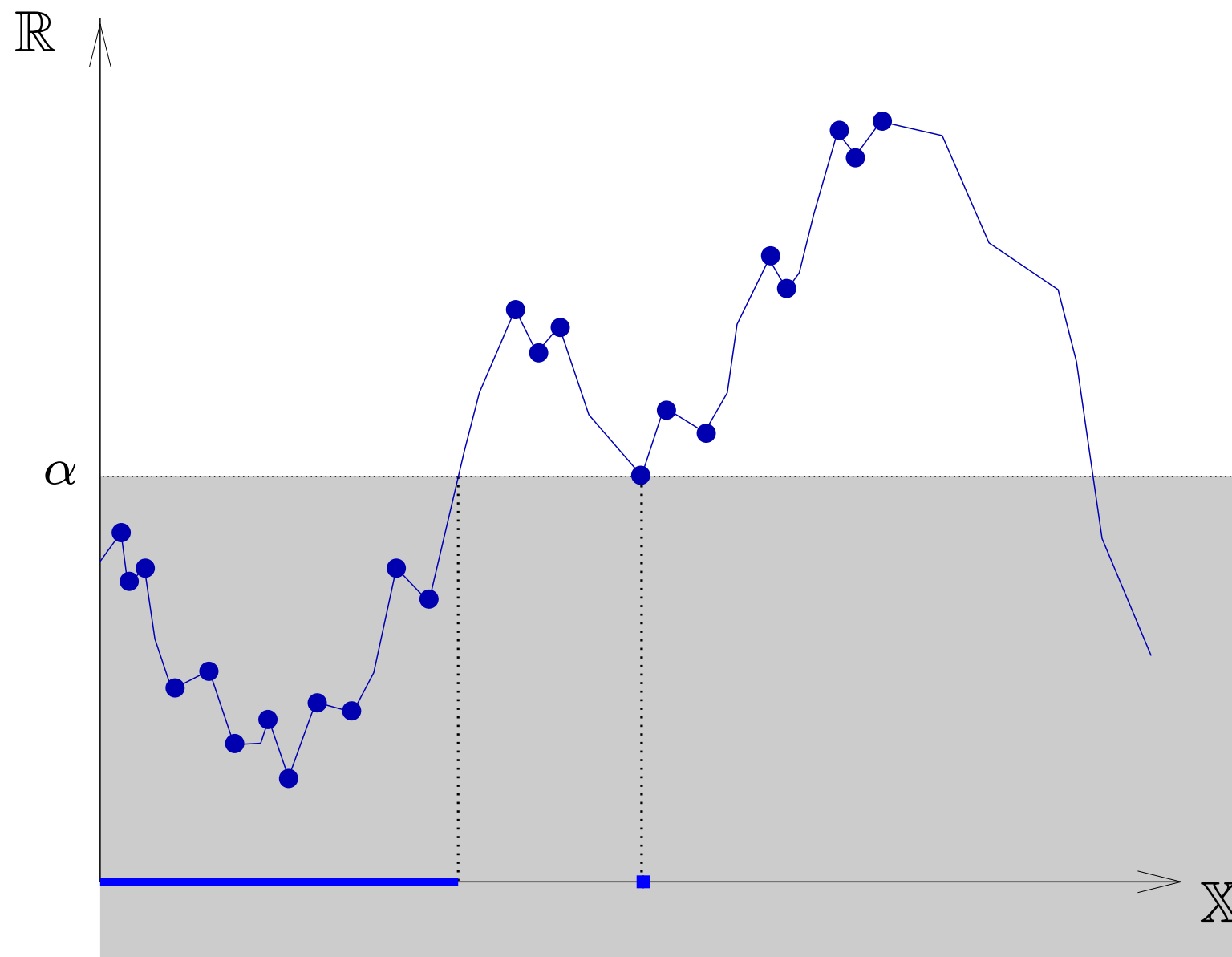
Persistence-Based Approach in a nutshell...

- evolution of topology of sub-level sets $\hat{f}^{-1}((-\infty, \alpha])$ as α spans $\mathbb{R}$.



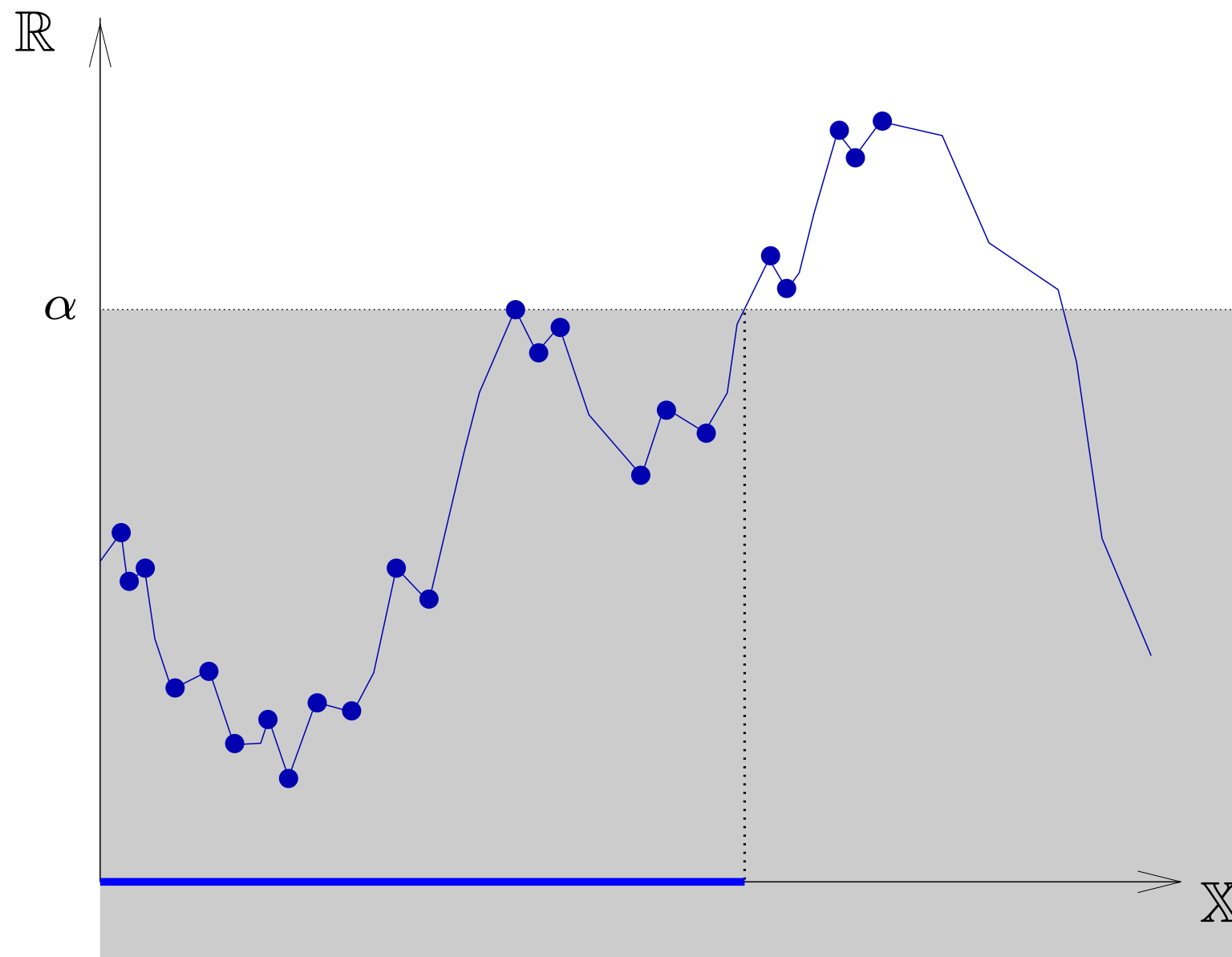
Persistence-Based Approach in a nutshell...

- evolution of topology of sub-level sets $\hat{f}^{-1}((-\infty, \alpha])$ as α spans $\mathbb{R}$.



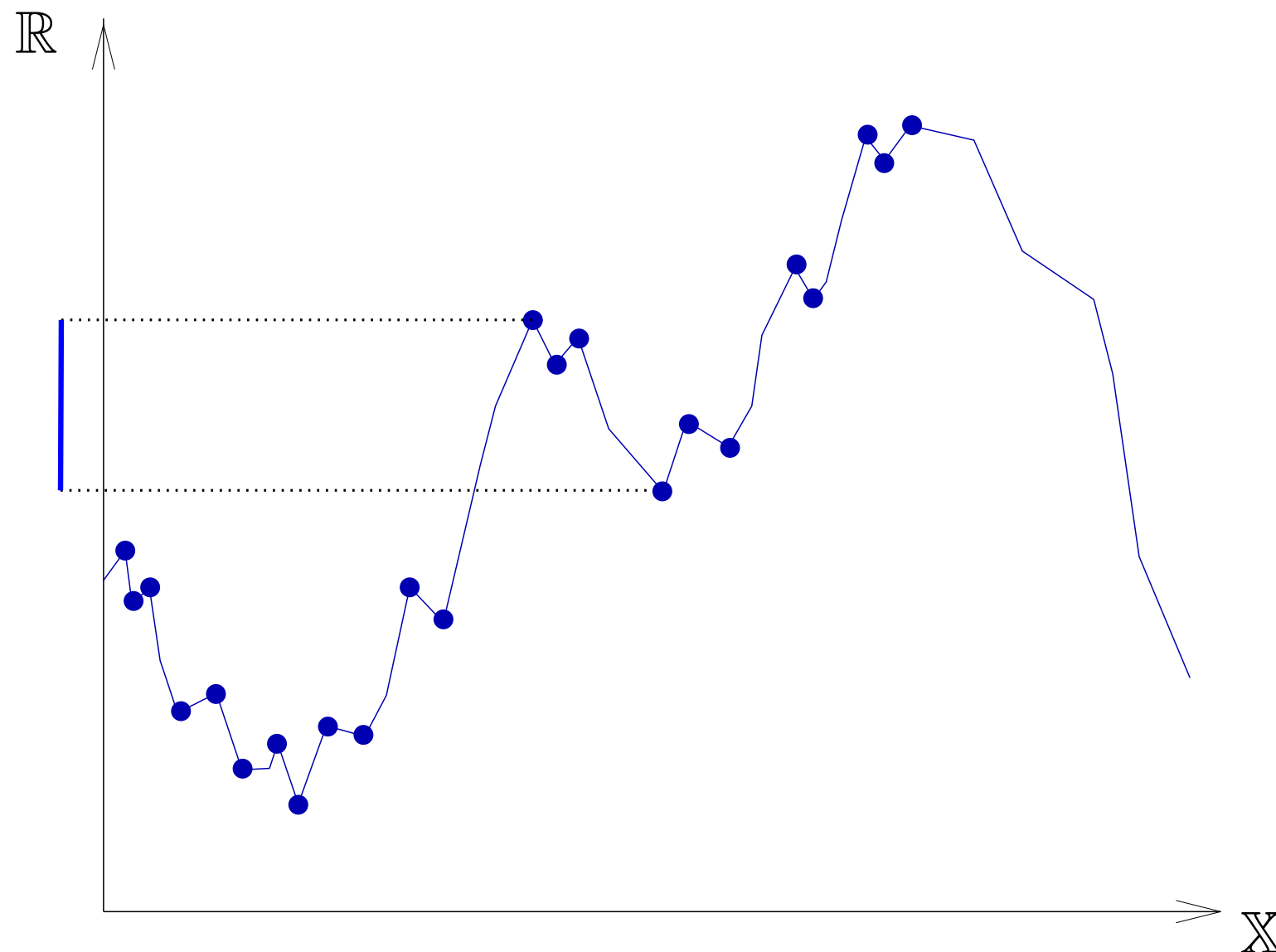
Persistence-Based Approach in a nutshell...

- evolution of topology of sub-level sets $\hat{f}^{-1}((-\infty, \alpha])$ as α spans $\mathbb{R}$.



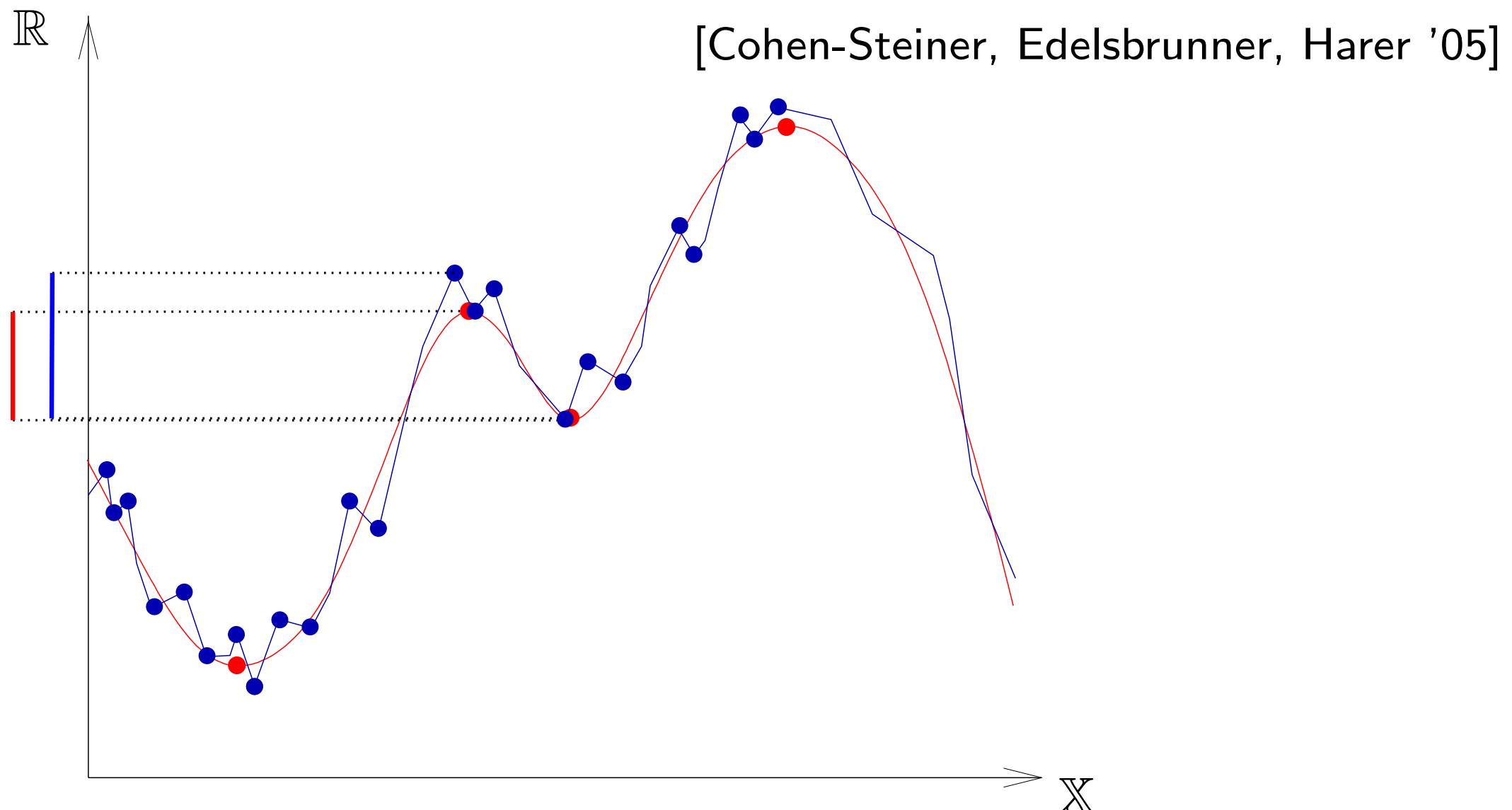
Persistence-Based Approach in a nutshell...

- evolution of topology of sub-level sets $\hat{f}^{-1}((-\infty, \alpha])$ as α spans $\mathbb{R}$.
- finite set of intervals (barcode) encode birth/death of homological features.



Persistence-Based Approach in a nutshell...

- evolution of topology of sub-level sets $\hat{f}^{-1}((-\infty, \alpha])$ as α spans $\mathbb{R}$.
- finite set of intervals (barcode) encode birth/death of homological features.
- barcode of $\hat{f}$ is close to barcode of f provided that $\|\hat{f} - f\|_\infty$ is small.

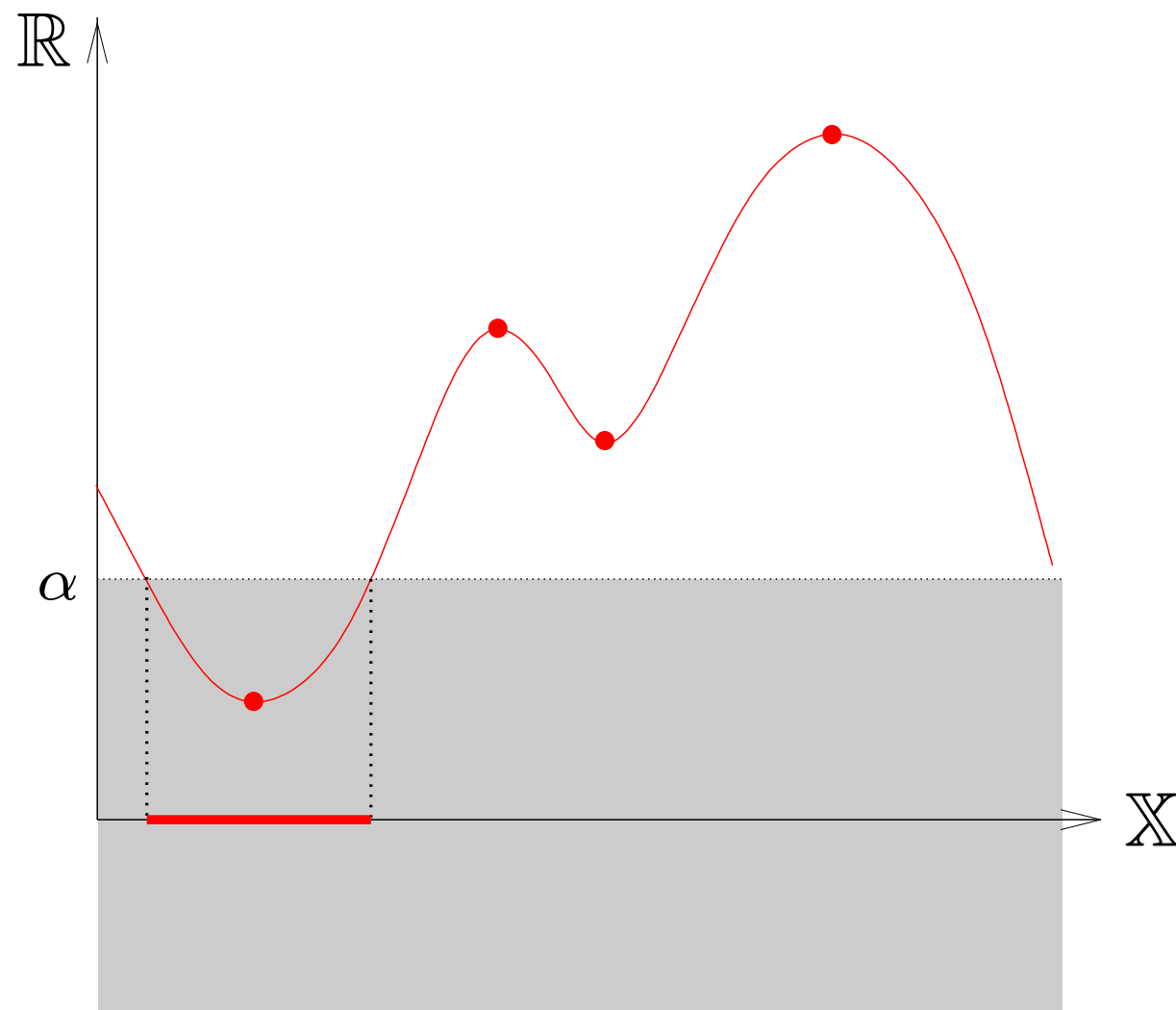


→ what if no triangulation of X is available?

- no geographic information,
- high-dimensional spaces,
- intrinsically curved spaces,
- non-triangulable spaces.

→ what if no triangulation of X is available?

- no geographic information,
- high-dimensional spaces,
- intrinsically curved spaces,
- non-triangulable spaces.



→ Approximate sub-level sets of f by other means

Approximation of Sub-Level Sets

- ~~Triangulation~~

Approximation of Sub-Level Sets

- ~~Triangulation~~
- Stability of persistence diagrams
 - interleaved spaces have close persistence diagrams

$$\begin{array}{ccccccc} \cdots & \longrightarrow & \phi_{a+2n\varepsilon} & \longrightarrow & \phi_{a+(2n+1)\varepsilon} & \longrightarrow & \phi_{a+(2n+2)\varepsilon} \longrightarrow \cdots \\ & \nearrow & & \searrow & & \nearrow & \searrow \\ \cdots & \longrightarrow & \psi_{a+2n\varepsilon} & \longrightarrow & \psi_{a+(2n+1)\varepsilon} & \longrightarrow & \psi_{a+(2n+2)\varepsilon} \longrightarrow \cdots \end{array}$$

Approximation of Sub-Level Sets

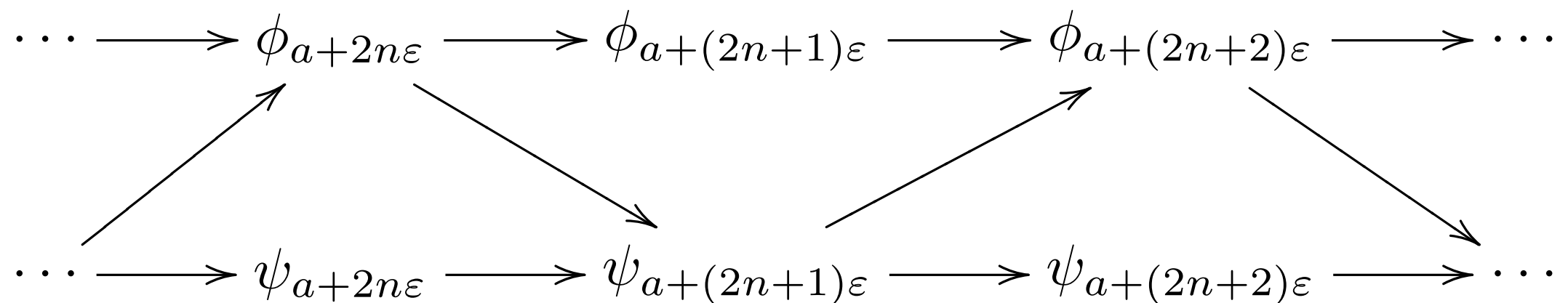
- ~~Triangulation~~
- Stability of persistence diagrams
 - interleaved spaces have close persistence diagrams

$$\begin{array}{ccccccc} \cdots & \longrightarrow & \phi_{a+2n\varepsilon} & \longrightarrow & \phi_{a+(2n+1)\varepsilon} & \longrightarrow & \phi_{a+(2n+2)\varepsilon} \longrightarrow \cdots \\ & \nearrow & & \searrow & & \nearrow & \\ \cdots & \longrightarrow & \psi_{a+2n\varepsilon} & \longrightarrow & \psi_{a+(2n+1)\varepsilon} & \longrightarrow & \psi_{a+(2n+2)\varepsilon} \longrightarrow \cdots \end{array}$$

- **Goal:** Relate level sets to something computable

Approximation of Sub-Level Sets

- ~~Triangulation~~
- Stability of persistence diagrams
 - interleaved spaces have close persistence diagrams



- **Goal:** Relate level sets to something computable

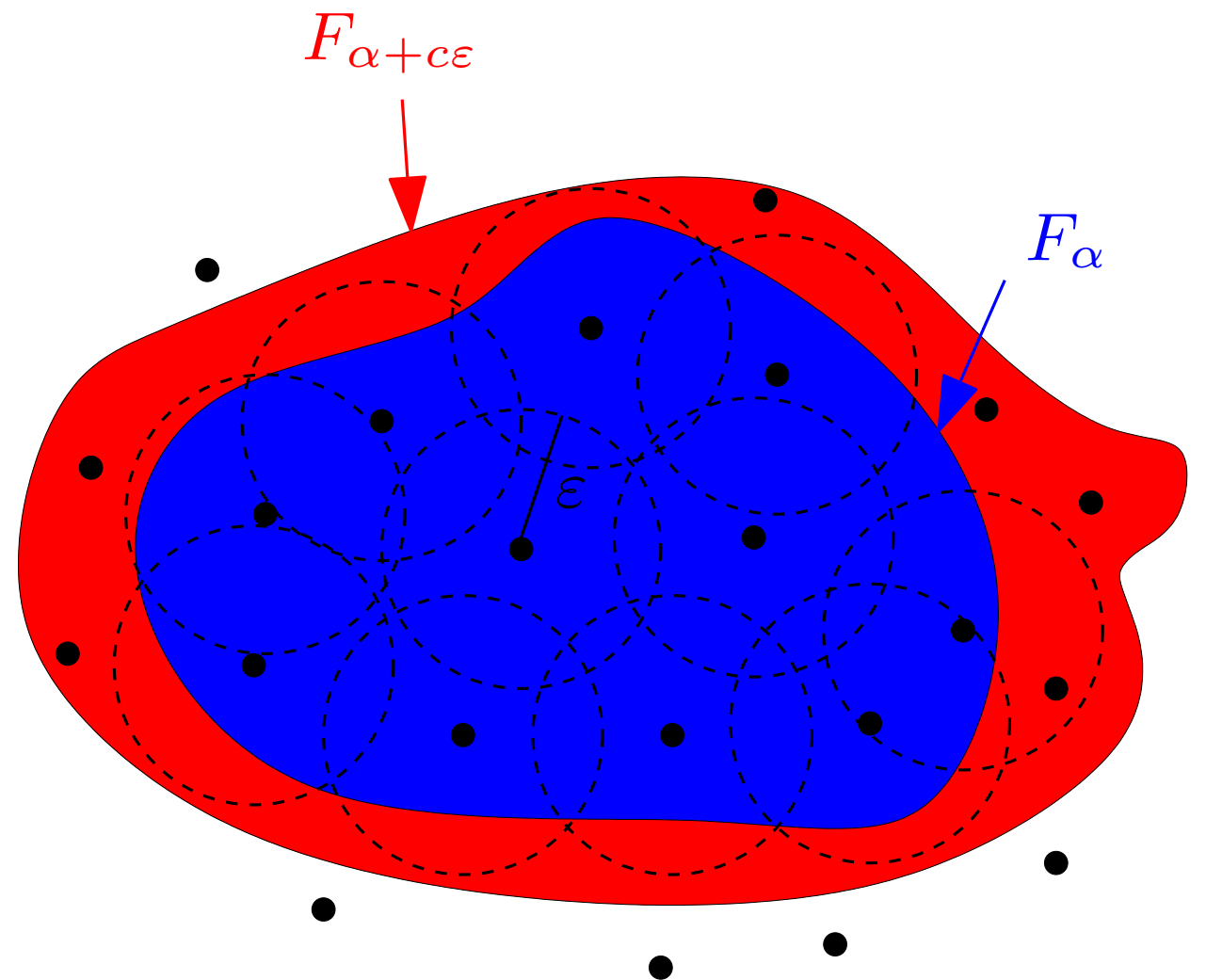
First step

Unions of Geodesic Balls

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

$$\left| \begin{array}{l} F_\alpha := f^{-1}((-\infty, \alpha]) \\ L_\alpha := L \cap F_\alpha \\ L_\alpha^\varepsilon := \bigcup_{p \in L_\alpha} B_{\mathbb{X}}(p, \varepsilon) \end{array} \right.$$

$$\boxed{\forall \alpha \in \mathbb{R}, F_\alpha \subseteq L_{\alpha+c\varepsilon}^\varepsilon \subseteq F_{\alpha+2c\varepsilon}}$$



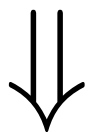
Unions of Geodesic Balls

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

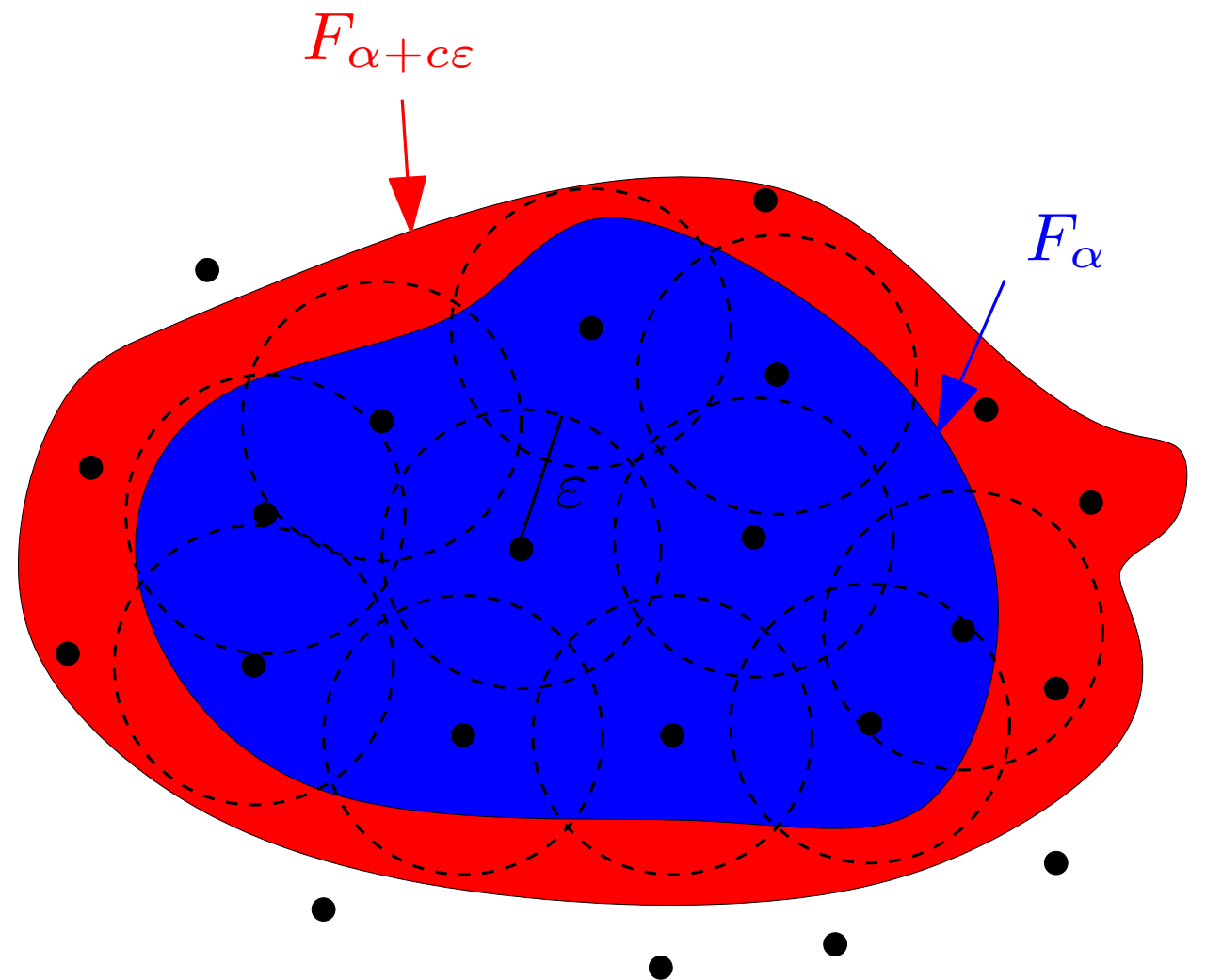
$$\left| \begin{array}{l} F_\alpha := f^{-1}((-\infty, \alpha]) \\ L_\alpha := L \cap F_\alpha \\ L_\alpha^\varepsilon := \bigcup_{p \in L_\alpha} B_{\mathbb{X}}(p, \varepsilon) \end{array} \right.$$

$$\boxed{\forall \alpha \in \mathbb{R}, F_\alpha \subseteq L_{\alpha+c\varepsilon}^\varepsilon \subseteq F_{\alpha+2c\varepsilon}}$$

the nested families of spaces
 $\{F_\alpha\}_{\alpha \in \mathbb{R}}$ and $\{L_\alpha^\varepsilon\}_{\alpha \in \mathbb{R}}$
 are $c\varepsilon$ -interleaved



their barcodes are $c\varepsilon$ -close.



Unions of Geodesic Balls

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

Algorithm: (parameter $\delta \geq 0$)

1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \dots \leq f(p_n)$,
2. For $i = 1, \dots, n$, build the union of balls $\{p_1, \dots, p_i\}^\delta$,
3. Apply the persistence algorithm to the nested family of spaces.

$$\{p_1\}^\delta \hookrightarrow \{p_1, p_2\}^\delta \hookrightarrow \{p_1, p_2, p_3\}^\delta \hookrightarrow \dots \hookrightarrow \{p_1, p_2, \dots, p_n\}^\delta$$

Unions of Geodesic Balls

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

Algorithm: (parameter $\delta \geq 0$)

1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \dots \leq f(p_n)$,
2. For $i = 1, \dots, n$, build the union of balls $\{p_1, \dots, p_i\}^\delta$,
or equivalently its nerve $\mathcal{N}^\delta(p_1, \dots, p_i)$,
3. Apply the persistence algorithm to the nested family of spaces.

$$\mathcal{N}^\delta(p_1) \hookrightarrow \mathcal{N}^\delta(p_1, p_2) \hookrightarrow \mathcal{N}^\delta(p_1, p_2, p_3) \hookrightarrow \dots \hookrightarrow \mathcal{N}^\delta(p_1, p_2, \dots, p_n)$$

Unions of Geodesic Balls

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

Algorithm: (parameter $\delta \geq 0$)

1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \dots \leq f(p_n)$,
2. For $i = 1, \dots, n$, build the union of balls $\{p_1, \dots, p_i\}^\delta$,
or equivalently its nerve $\mathcal{N}^\delta(p_1, \dots, p_i)$,
3. Apply the persistence algorithm to the nested family of spaces.

$$\mathcal{N}^\delta(p_1) \hookrightarrow \mathcal{N}^\delta(p_1, p_2) \hookrightarrow \mathcal{N}^\delta(p_1, p_2, p_3) \hookrightarrow \dots \hookrightarrow \mathcal{N}^\delta(p_1, p_2, \dots, p_n)$$

$\rightarrow$ if $\delta \geq \varepsilon$, get a $c\delta$ -approximation of the barcode of f

Unions of Geodesic Balls

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ continuous, L geodesic ε -cover of $\mathbb{X}$, for some $\varepsilon > 0$.

Algorithm: (parameter $\delta \geq 0$)

1. Sort the data points such that $f(p_1) \leq \dots \leq f(p_n)$,
2. For $i = 1, \dots, n$, build the union $U_i = \{p_1, \dots, p_i\}^\delta$,
or equivalently nerve $\mathcal{N}^\delta(p_1, \dots, p_i)$,
3. Apply the persistence barcode to the nested family of spaces.

$$\mathcal{N}^\delta(p_1) \hookrightarrow \mathcal{N}^\delta(p_1, p_2) \hookrightarrow \mathcal{N}^\delta(p_1, p_2, p_3) \hookrightarrow \dots \hookrightarrow \mathcal{N}^\delta(p_1, p_2, \dots, p_n)$$

Pairs of Rips Complexes

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

→ idea inspired from [Chazal, Oudot '08]:

$\forall i = 1, \dots, n$, replace $\{p_1, \dots, p_i\}^\delta$ by $\mathcal{R}^\delta(p_1, \dots, p_i) \subseteq \mathcal{R}^{2\delta}(p_1, \dots, p_i)$

Pairs of Rips Complexes

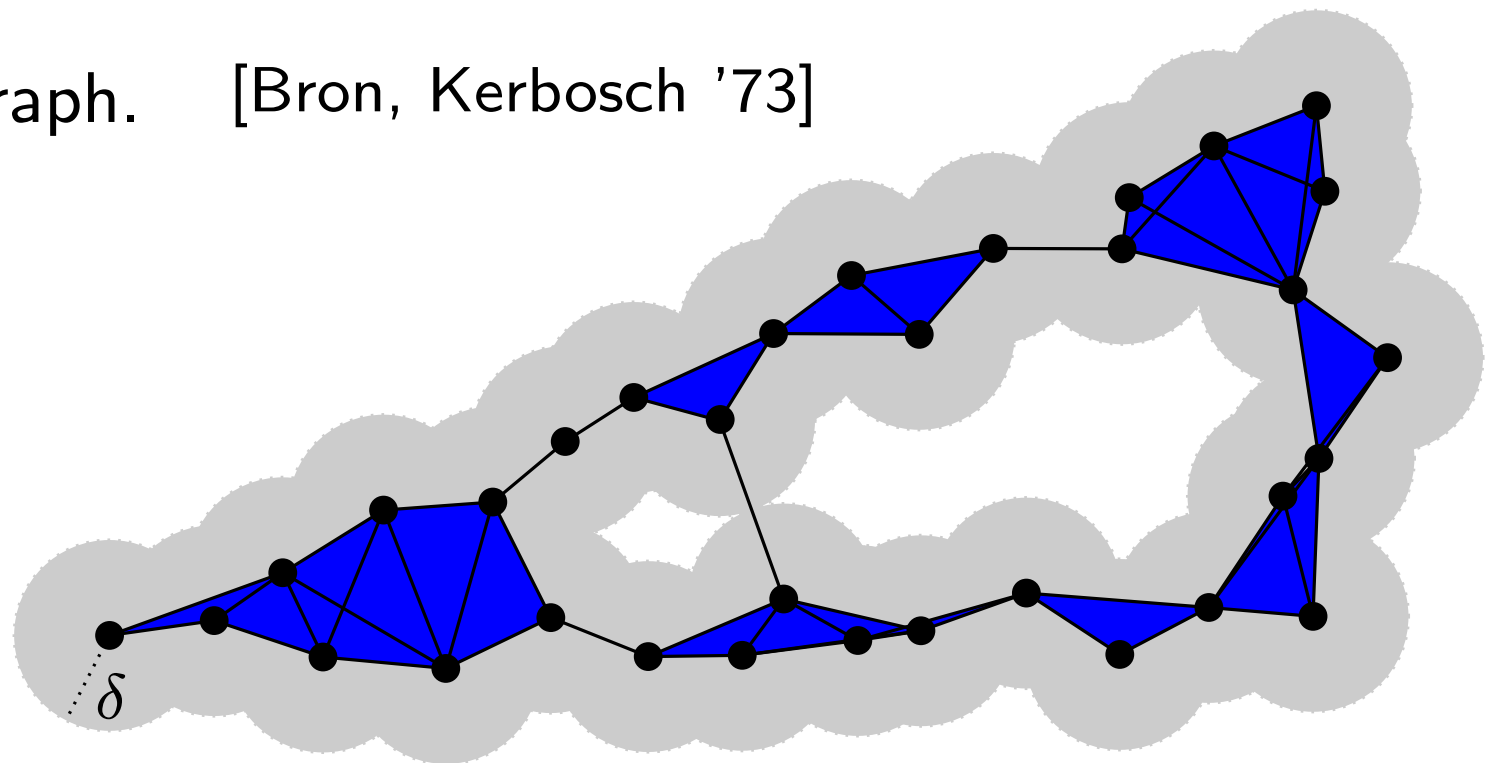
Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

→ idea inspired from [Chazal, Oudot '08]:

$\forall i = 1, \dots, n$, replace $\{p_1, \dots, p_i\}^\delta$ by $\mathcal{R}^\delta(p_1, \dots, p_i) \subseteq \mathcal{R}^{2\delta}(p_1, \dots, p_i)$

Construction of $\mathcal{R}^{2\delta}(p_1, \dots, p_i)$:

- connect every pair of points (p_j, p_k) such that $d_{\mathbb{X}}(p_j, p_k) \leq 2\delta$,
- report all cliques in the graph. [Bron, Kerbosch '73]



Pairs of Rips Complexes

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

→ idea inspired from [Chazal, Oudot '08]:

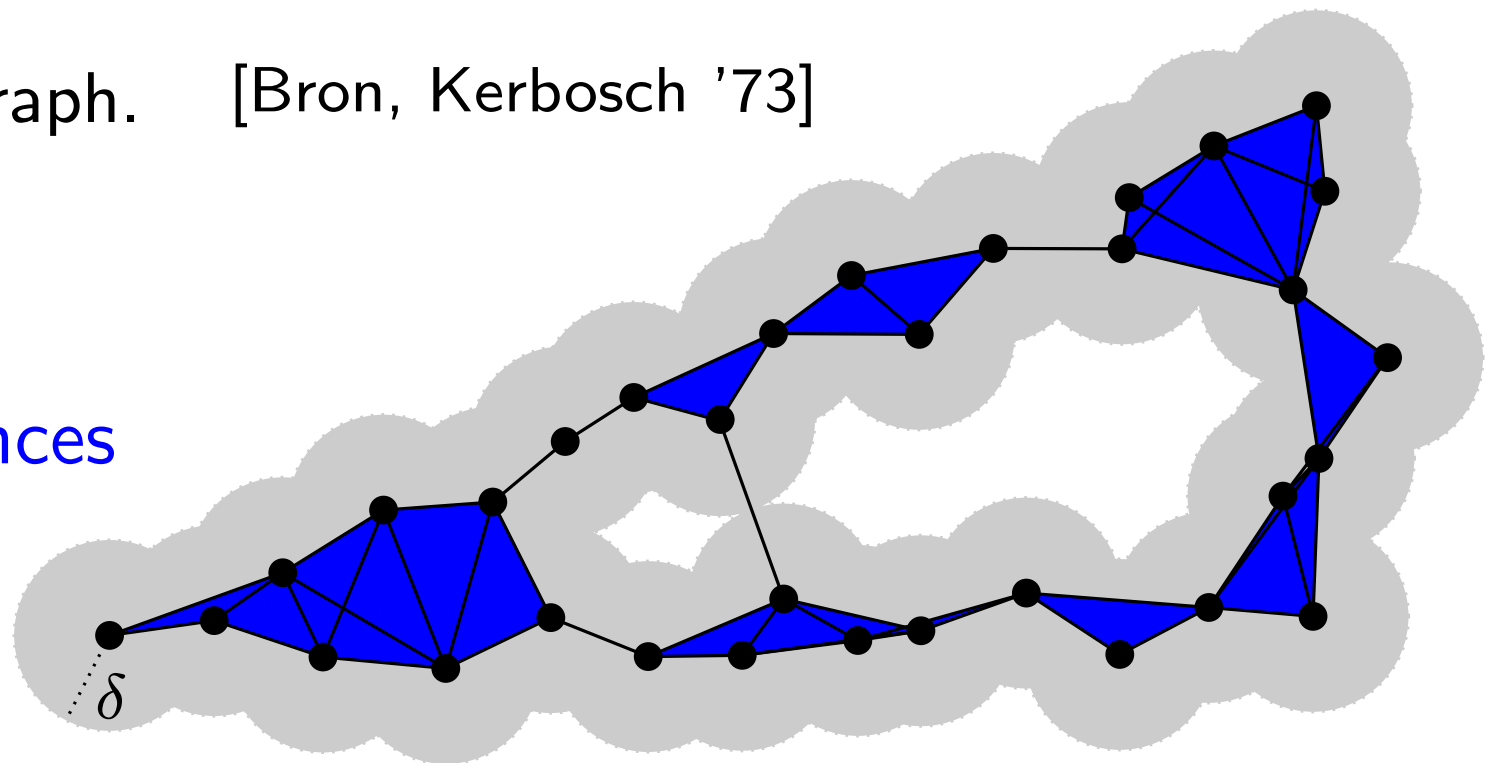
$\forall i = 1, \dots, n$, replace $\{p_1, \dots, p_i\}^\delta$ by $\mathcal{R}^\delta(p_1, \dots, p_i) \subseteq \mathcal{R}^{2\delta}(p_1, \dots, p_i)$

Construction of $\mathcal{R}^{2\delta}(p_1, \dots, p_i)$:

- connect every pair of points (p_j, p_k) such that $d_{\mathbb{X}}(p_j, p_k) \leq 2\delta$,
- report all cliques in the graph. [Bron, Kerbosch '73]

→ takes $O(n^2 + Cn)$ time

→ uses only geodesic distances
between the data points



Pairs of Rips Complexes

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

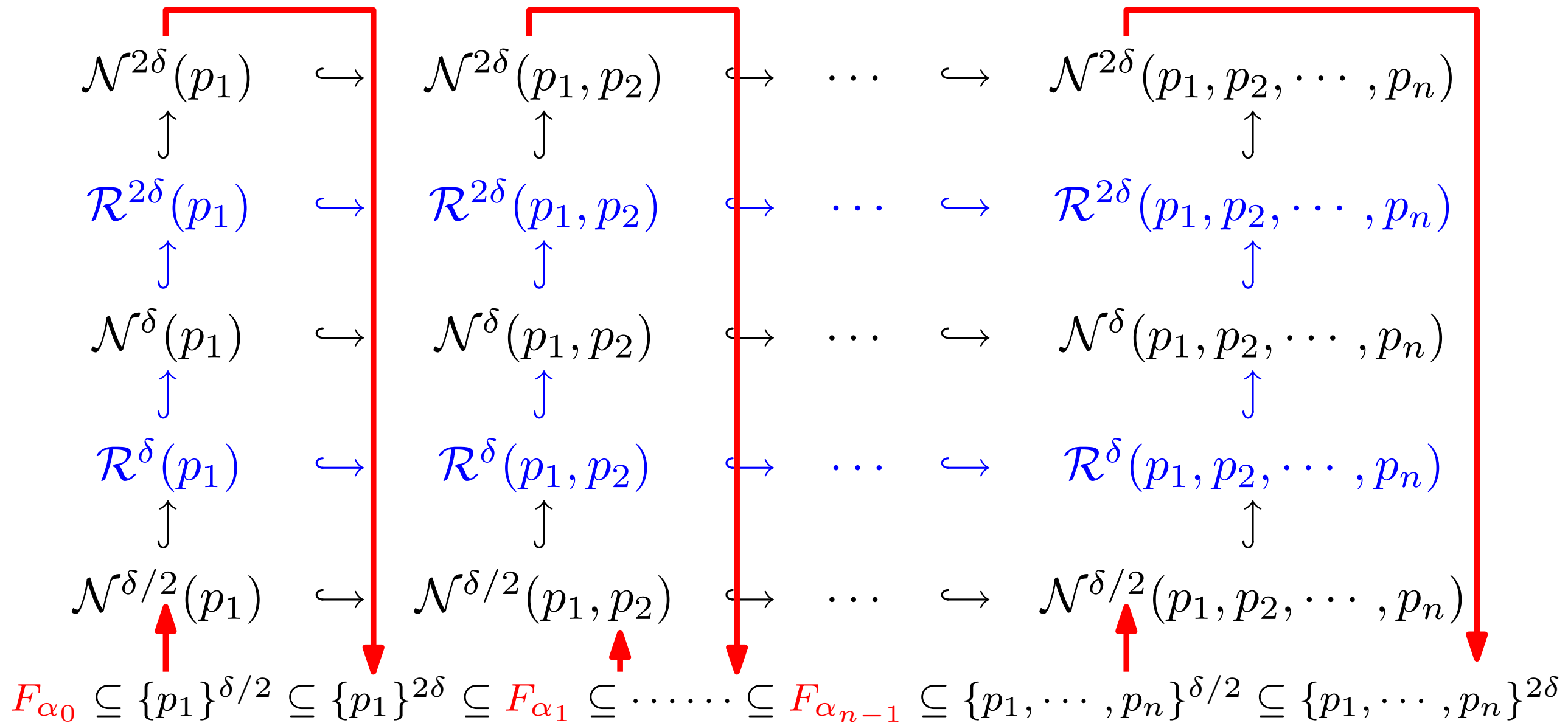
Guarantees:

$$\begin{array}{ccccccc}
 \mathcal{N}^{2\delta}(p_1) & \hookrightarrow & \mathcal{N}^{2\delta}(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{N}^{2\delta}(p_1, p_2, \dots, p_n) \\
 \uparrow & & \uparrow & & & & \uparrow \\
 \mathcal{R}^{2\delta}(p_1) & \hookrightarrow & \mathcal{R}^{2\delta}(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{R}^{2\delta}(p_1, p_2, \dots, p_n) \\
 \uparrow & & \uparrow & & & & \uparrow \\
 \mathcal{N}^{\delta}(p_1) & \hookrightarrow & \mathcal{N}^{\delta}(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{N}^{\delta}(p_1, p_2, \dots, p_n) \\
 \uparrow & & \uparrow & & & & \uparrow \\
 \mathcal{R}^{\delta}(p_1) & \hookrightarrow & \mathcal{R}^{\delta}(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{R}^{\delta}(p_1, p_2, \dots, p_n) \\
 \uparrow & & \uparrow & & & & \uparrow \\
 \mathcal{N}^{\delta/2}(p_1) & \hookrightarrow & \mathcal{N}^{\delta/2}(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{N}^{\delta/2}(p_1, p_2, \dots, p_n)
 \end{array}$$

Pairs of Rips Complexes

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

Guarantees:



Pairs of Rips Complexes

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

Guarantees:

$\forall \delta \geq \varepsilon$, $\{F_\alpha\}_{\alpha \in \mathbb{R}}$ and $\{\mathcal{R}^\delta(L_\alpha) \hookrightarrow \mathcal{R}^{2\delta}(L_\alpha)\}_{\alpha \in \mathbb{R}}$ are $2c\delta$ -interleaved

$\Downarrow$ [Chazal, Cohen-Steiner, Glisse, Guibas, Oudot '09]

their barcodes are $2c\delta$ -close.

Pairs of Rips Complexes

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

Algorithm: (parameter $\delta \geq 0$)

1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \dots \leq f(p_n)$,
2. For $i = 1, \dots, n$, build $\mathcal{R}^\delta(p_1, \dots, p_i) \subseteq \mathcal{R}^{2\delta}(p_1, \dots, p_i)$,
3. Apply the persistence algorithm **for images** to the pair of filtrations.

[Cohen-Steiner, Edelsbrunner, Harer, Morozov '09]

$$\begin{array}{ccccccc}
 \mathcal{R}^{2\delta}(p_1) & \hookrightarrow & \mathcal{R}^{2\delta}(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{R}^{2\delta}(p_1, p_2, \dots, p_n) \\
 \uparrow & & \uparrow & & & & \uparrow \\
 \mathcal{R}^\delta(p_1) & \hookrightarrow & \mathcal{R}^\delta(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{R}^\delta(p_1, p_2, \dots, p_n)
 \end{array}$$

Pairs of Rips Complexes

Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.

Algorithm: (parameter $\delta \geq 0$)

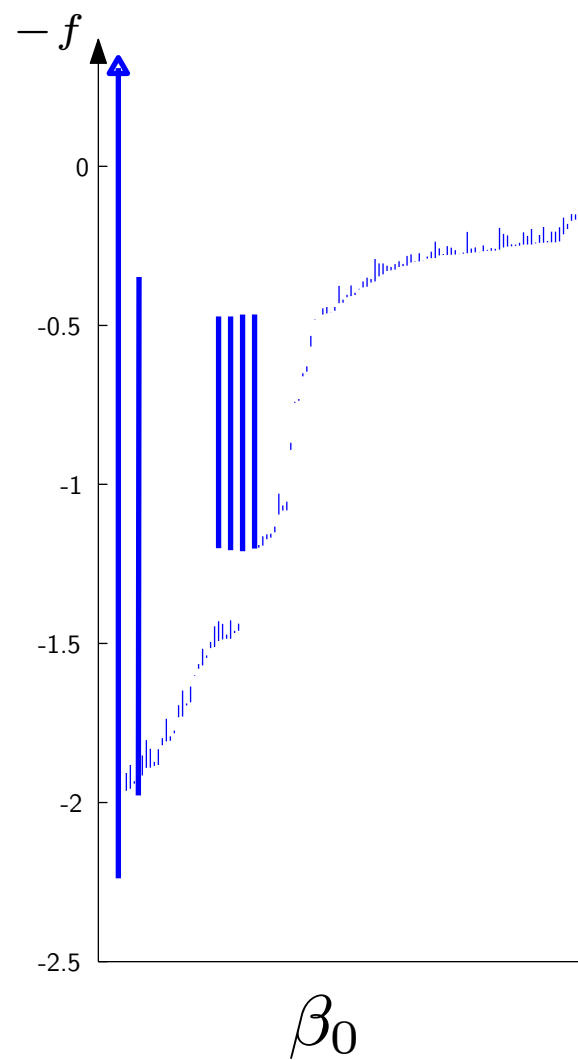
1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \dots \leq f(p_n)$,
2. For $i = 1, \dots, n$, build $\mathcal{R}^\delta(p_1, \dots, p_i) \subseteq \mathcal{R}^{2\delta}(p_1, \dots, p_i)$,
3. Apply the persistence algorithm **for images** to the pair of filtrations.

[Cohen-Steiner, Edelsbrunner, Harer, Morozov '09]

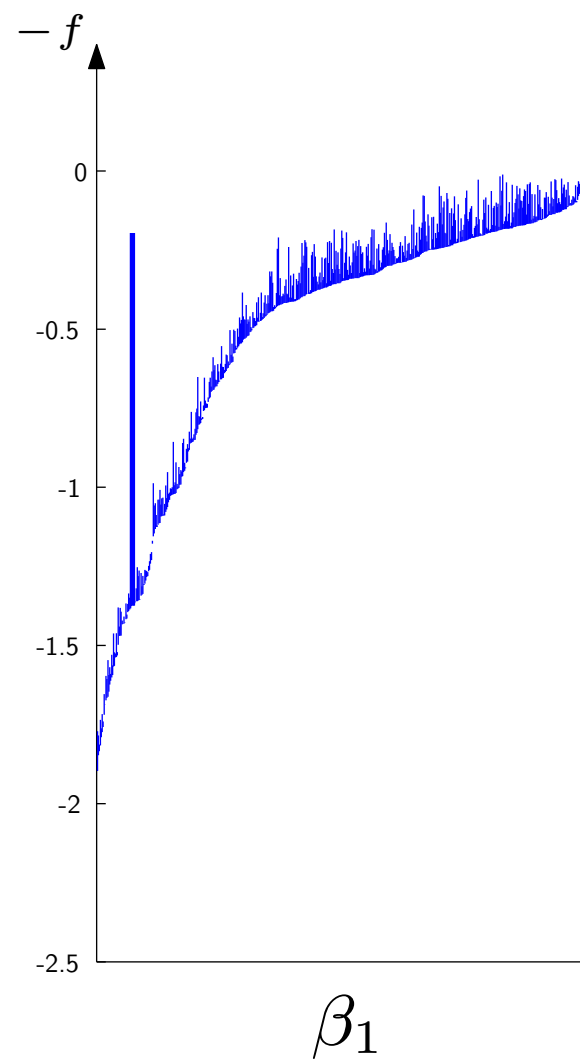
$$\begin{array}{ccccccc}
 \mathcal{R}^{2\delta}(p_1) & \hookrightarrow & \mathcal{R}^{2\delta}(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{R}^{2\delta}(p_1, p_2, \dots, p_n) \\
 \uparrow & & \uparrow & & & & \uparrow \\
 \mathcal{R}^\delta(p_1) & \hookrightarrow & \mathcal{R}^\delta(p_1, p_2) & \hookrightarrow & \dots & \hookrightarrow & \mathcal{R}^\delta(p_1, p_2, \dots, p_n)
 \end{array}$$

Pairs of Rips Complexes

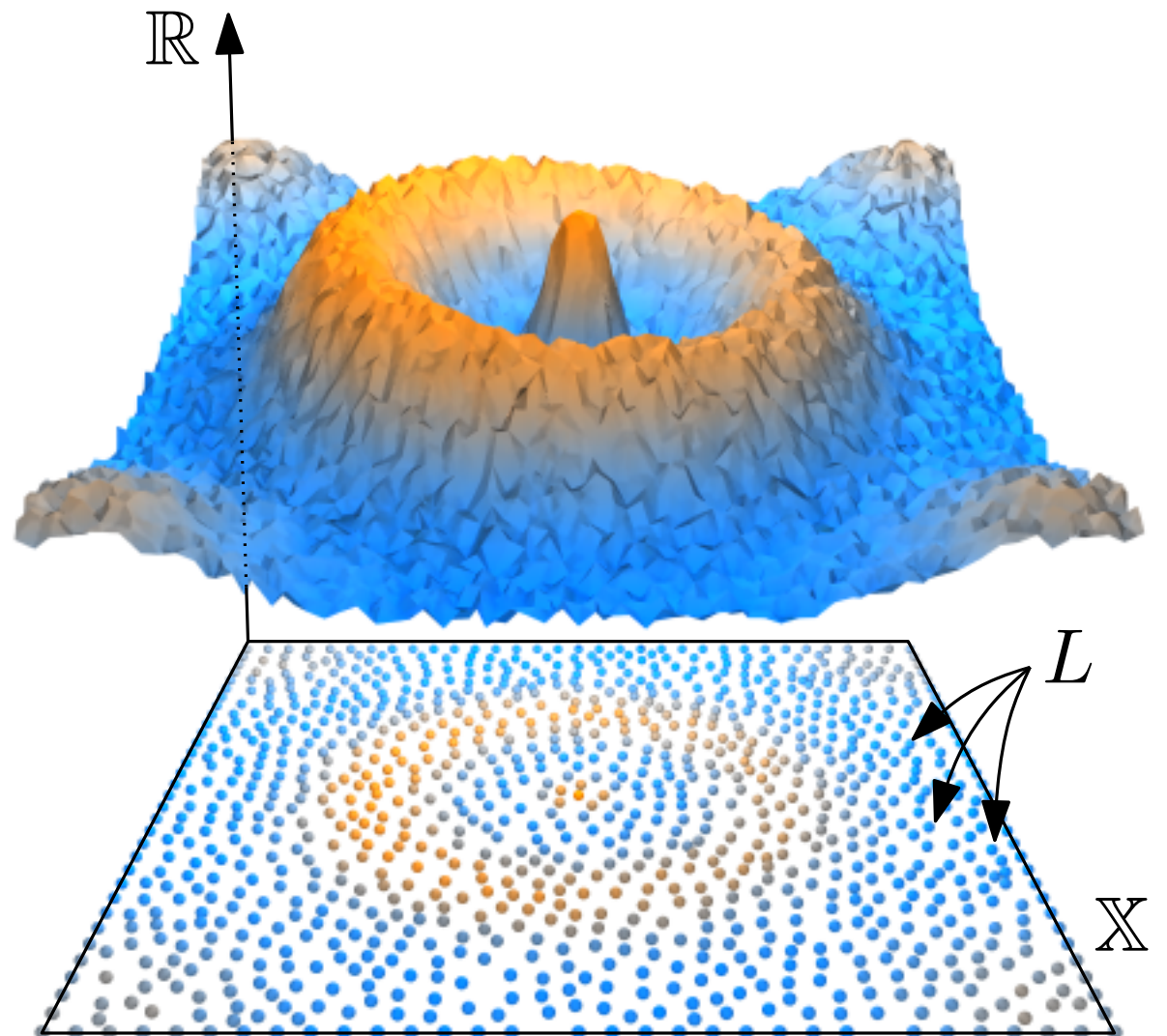
Assumptions: $\mathbb{X}$ Riemannian manifold, $f : \mathbb{X} \rightarrow \mathbb{R}$ c -Lipschitz,
 L geodesic ε -cover of $\mathbb{X}$, for some unknown $\varepsilon > 0$.



(6 prominent peaks)



(ring-shaped basin of attraction)



Basins of Attraction

Goal: approximate basins of attraction of significant peaks of f

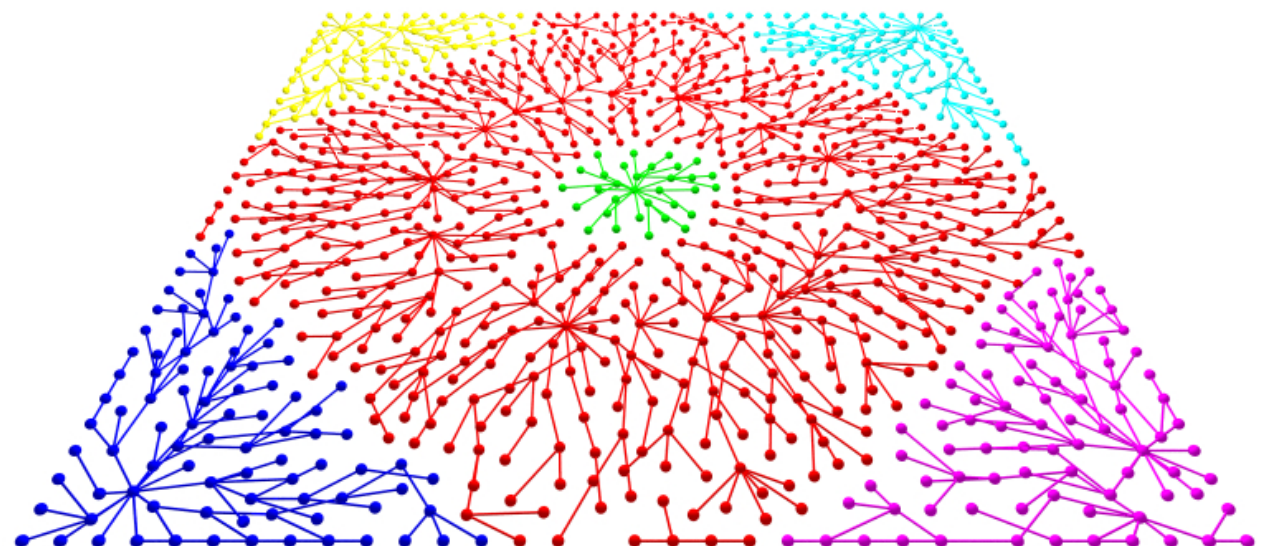
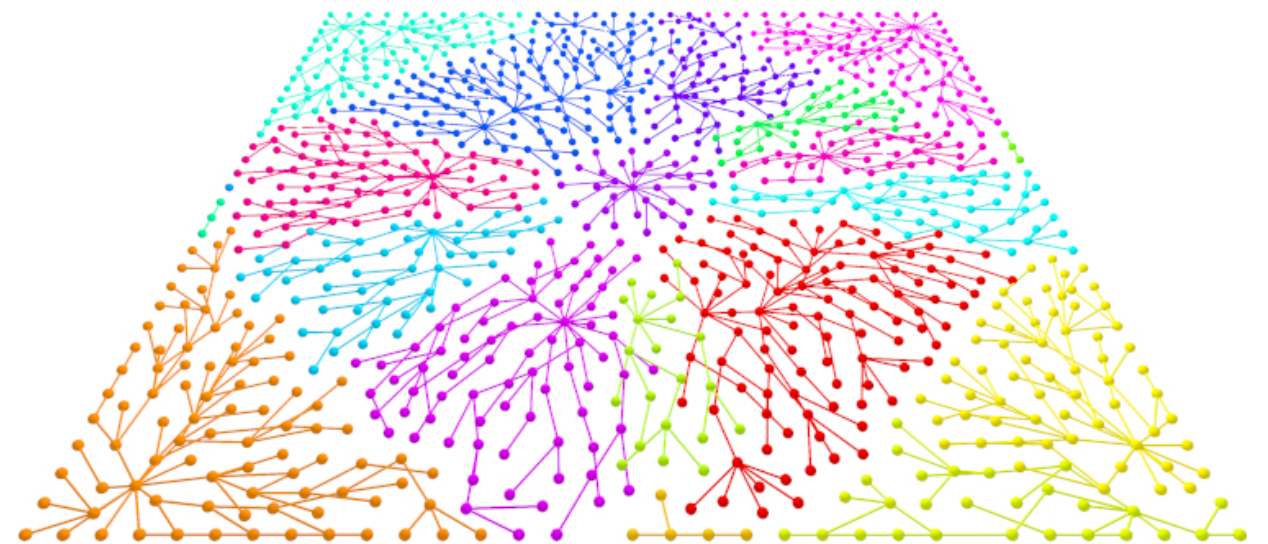
$\Rightarrow$ segmentation/clustering of point cloud L

Approach:

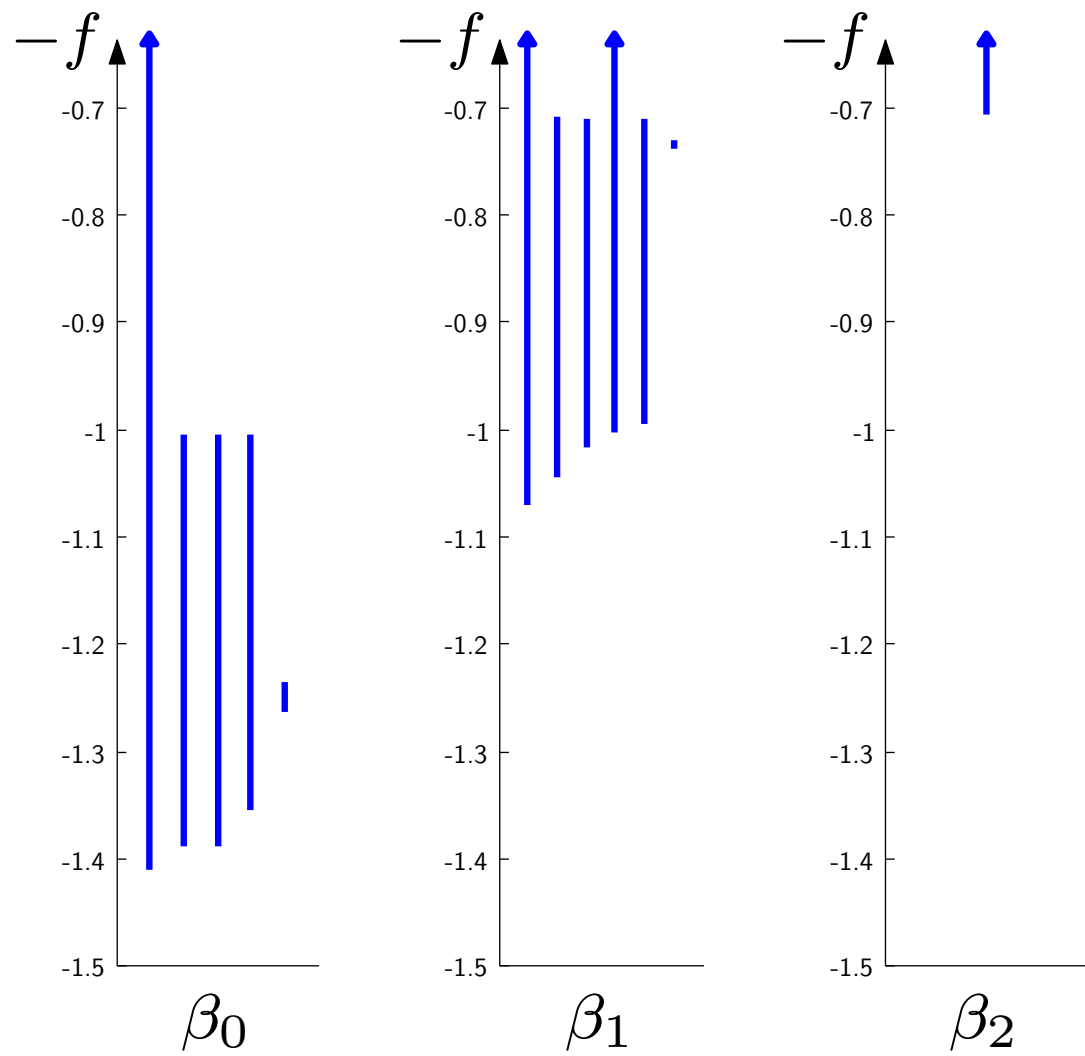
- rough approximation of gradient of f within Rips graph,
- merge clusters according to 0-dimensional barcode.

$\rightarrow$ union-find data structure

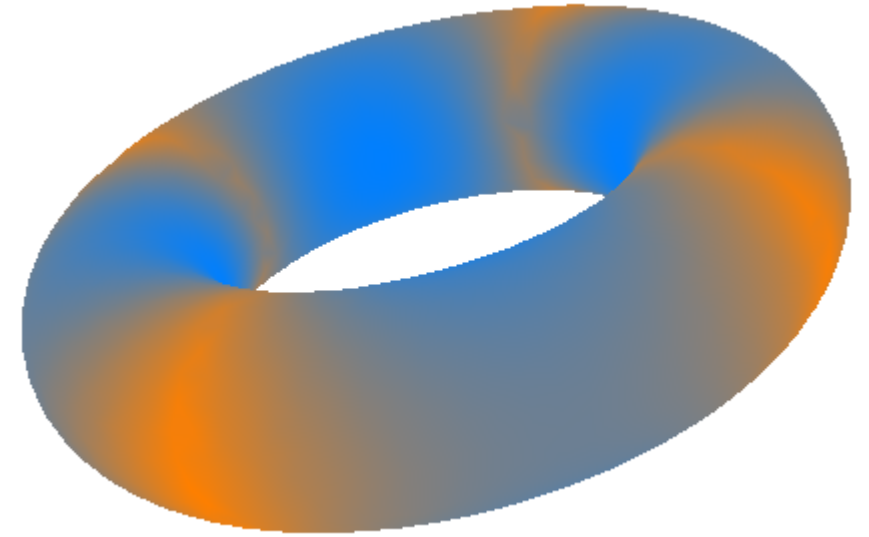
(see SODA '09
paper for details)



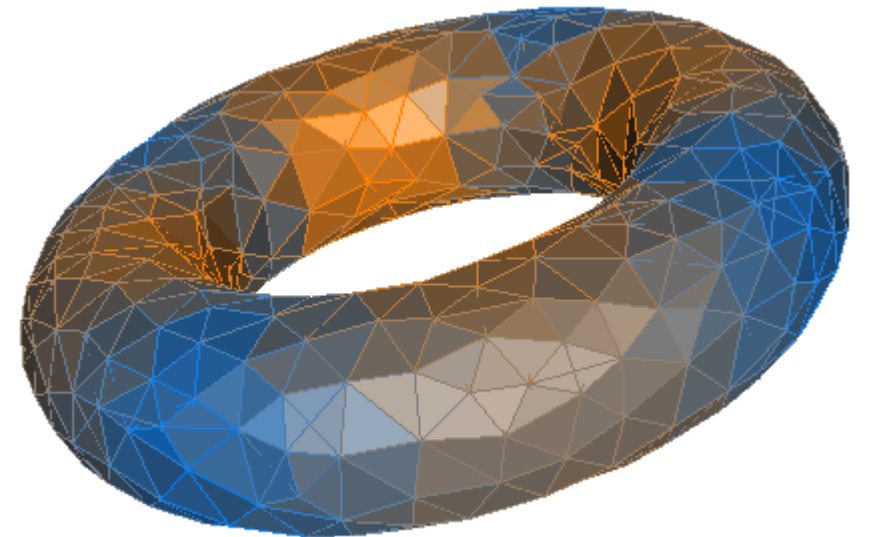
Some Results



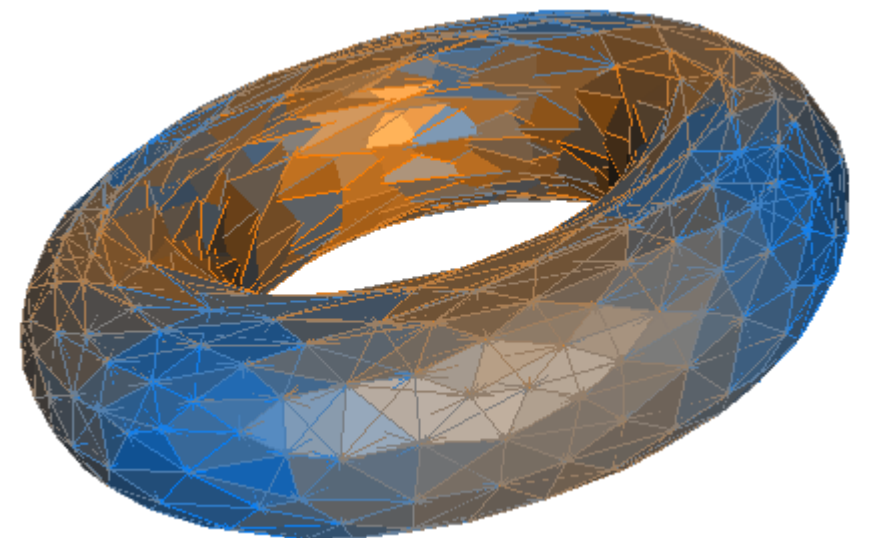
$(\mathbb{X}, f)$



$\mathcal{R}^\delta(L)$



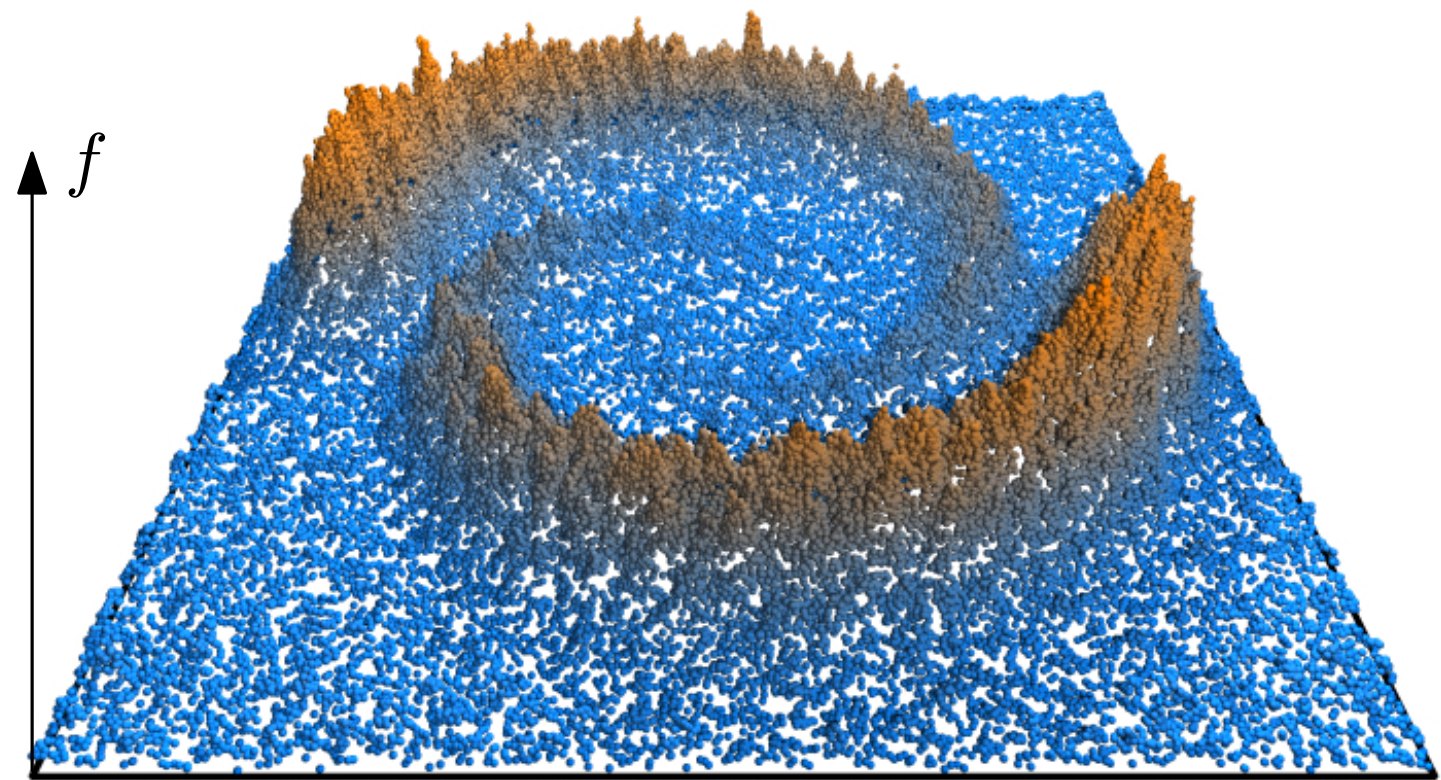
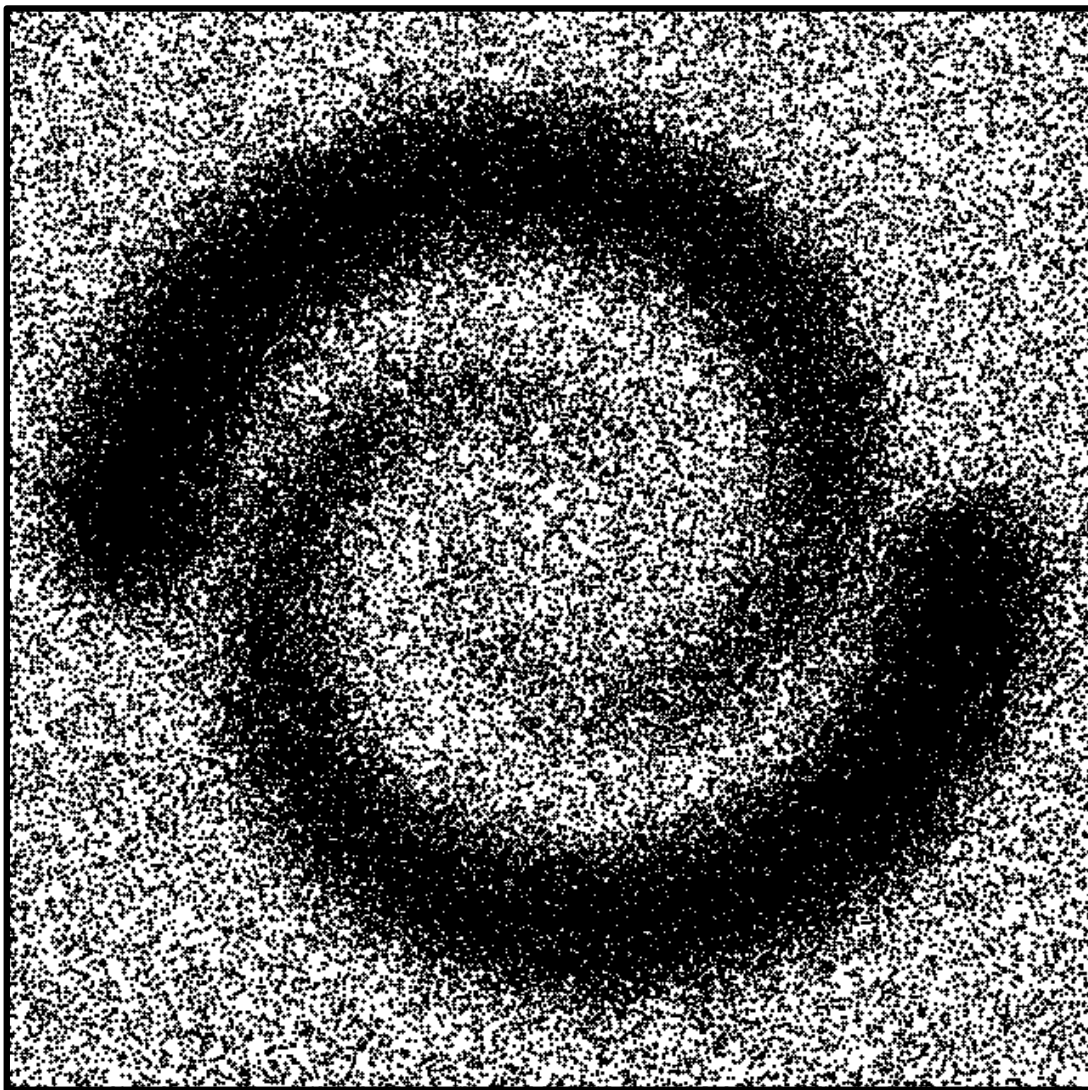
$\mathcal{R}^{2\delta}(L)$



Some Results

Input: $\mathbb{X} = [0, 1]^2$; $|L| = 100,000$;

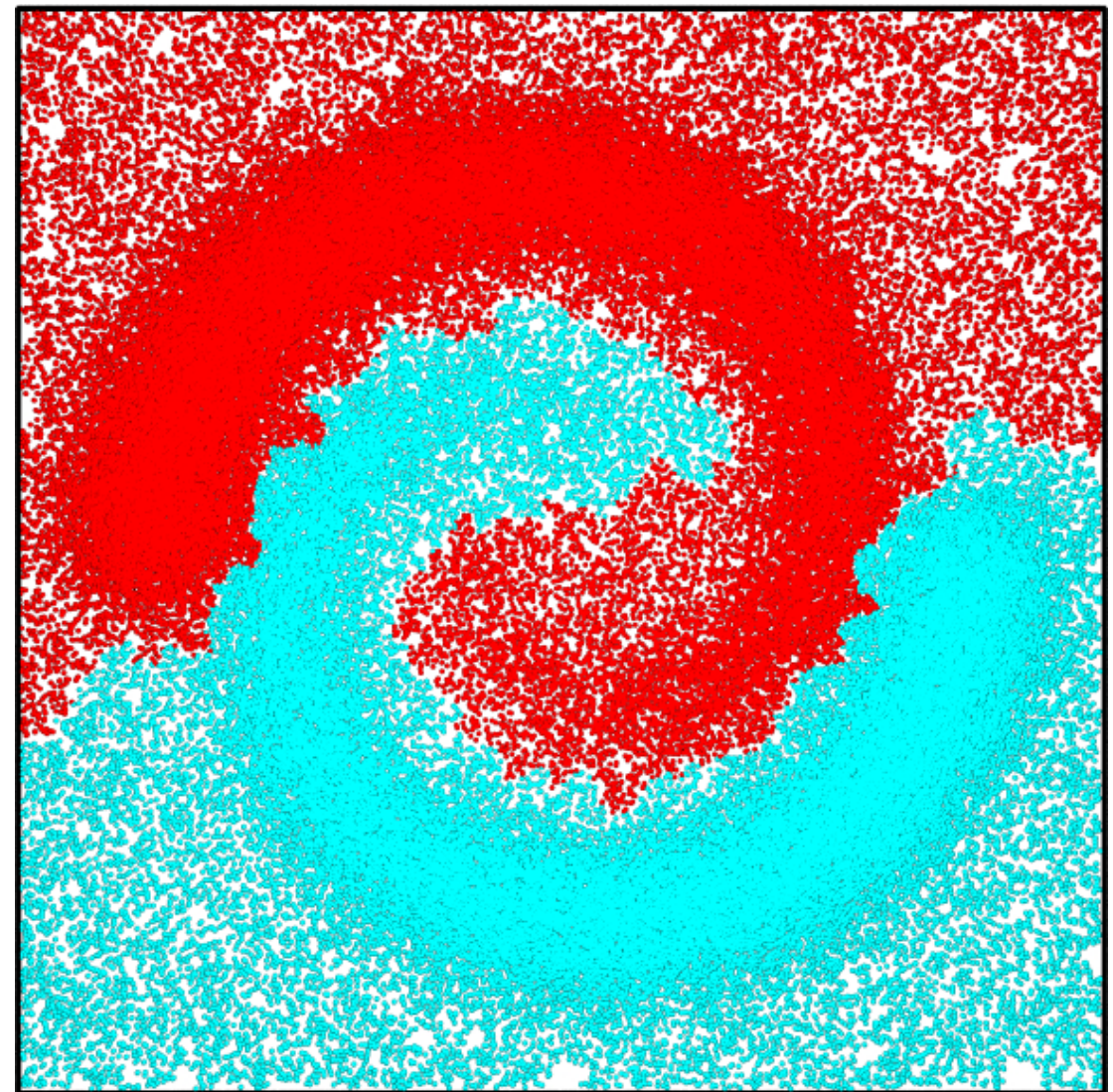
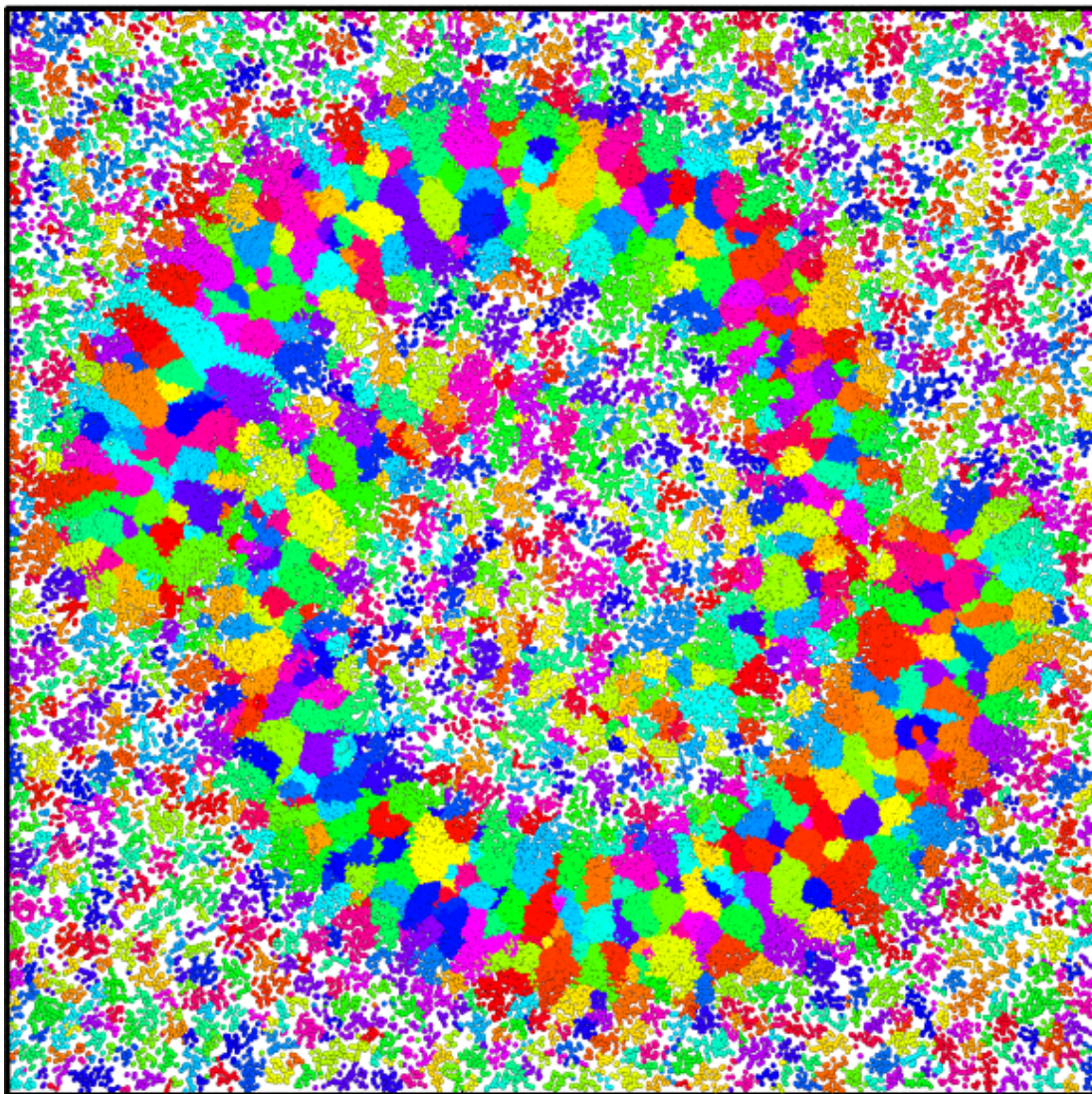
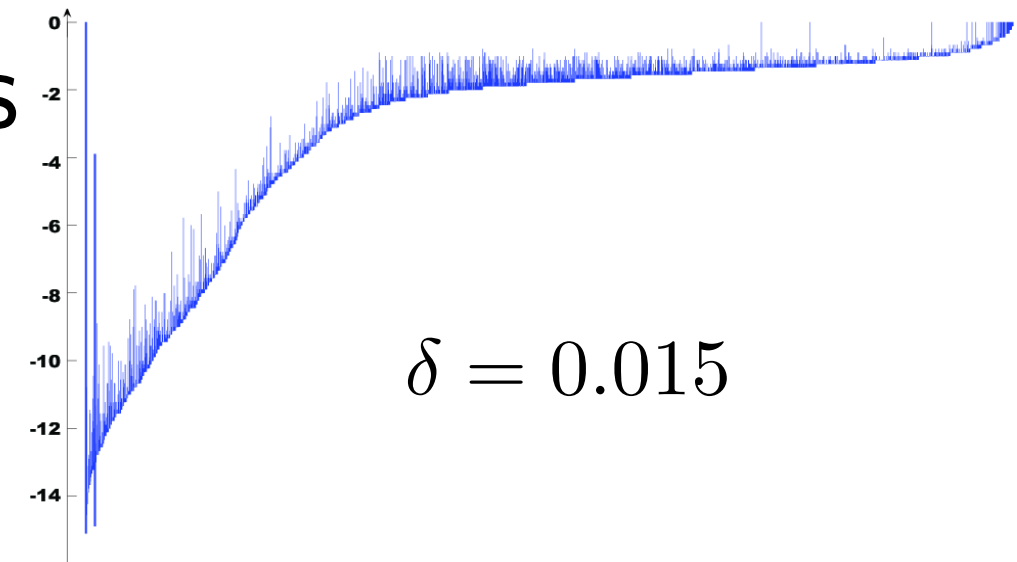
$f = \# \{ \text{data pts in fixed-radius ball} \}$



Some Results

Input: $\mathbb{X} = [0, 1]^2$; $|L| = 100,000$;

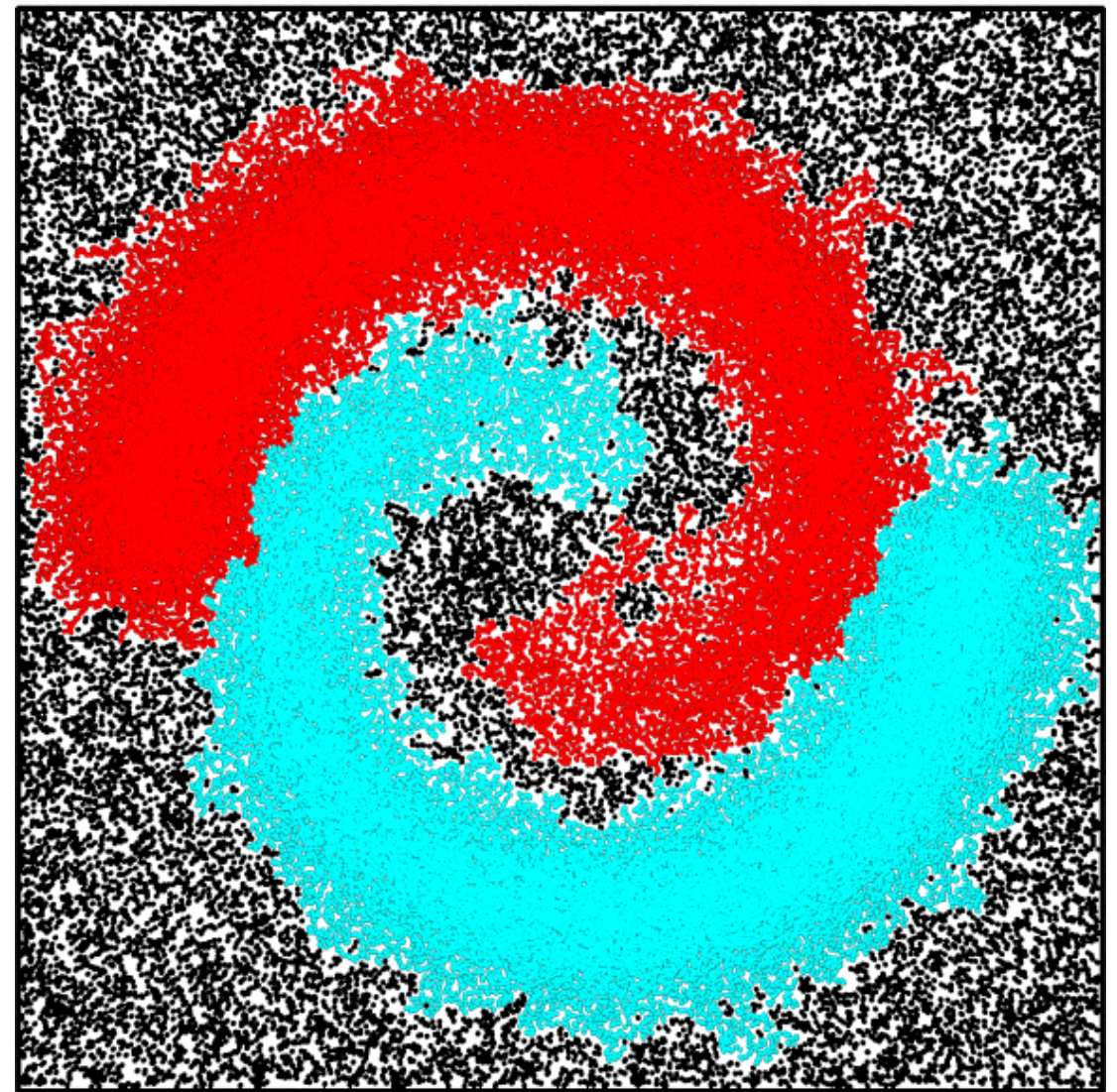
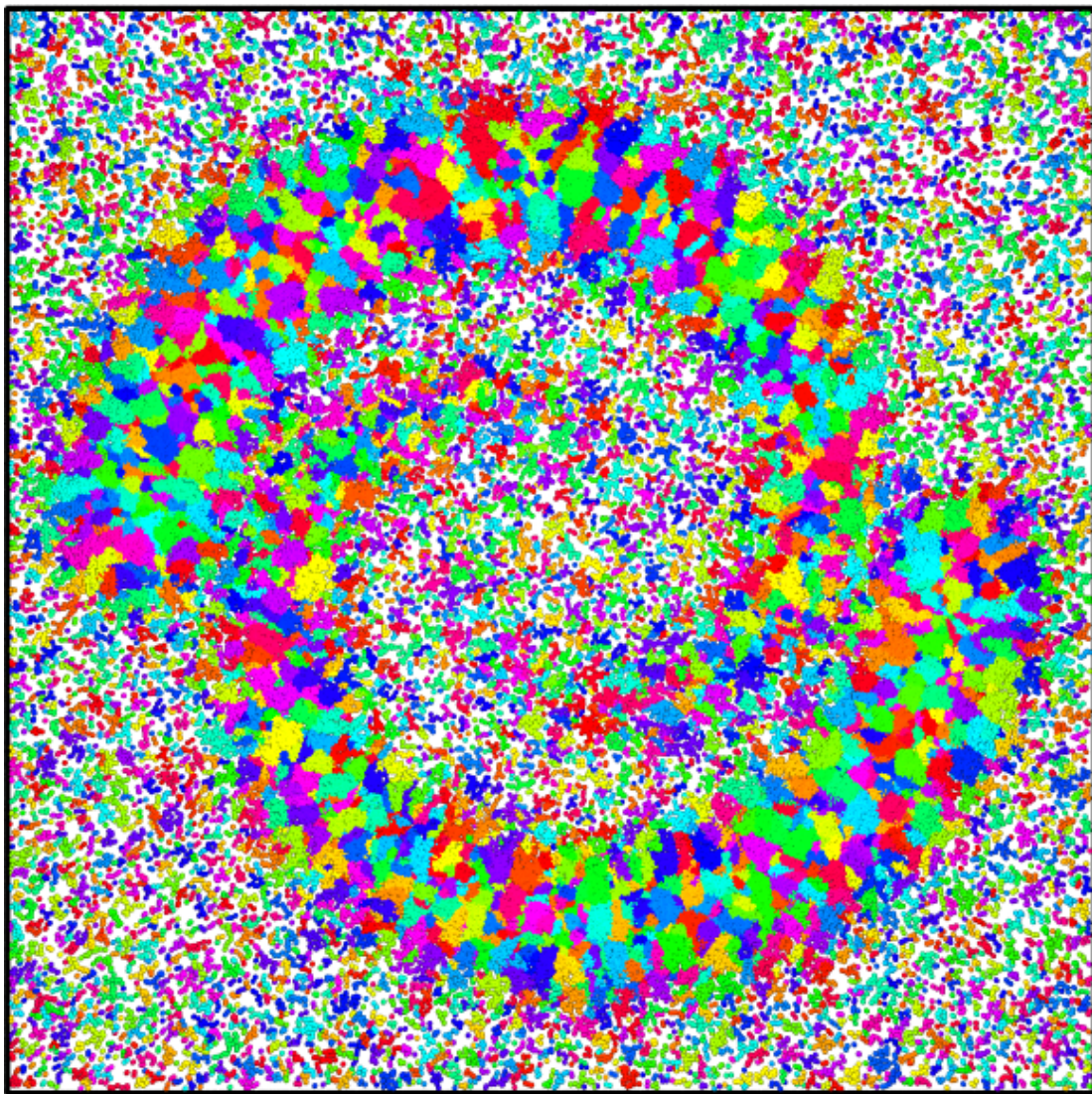
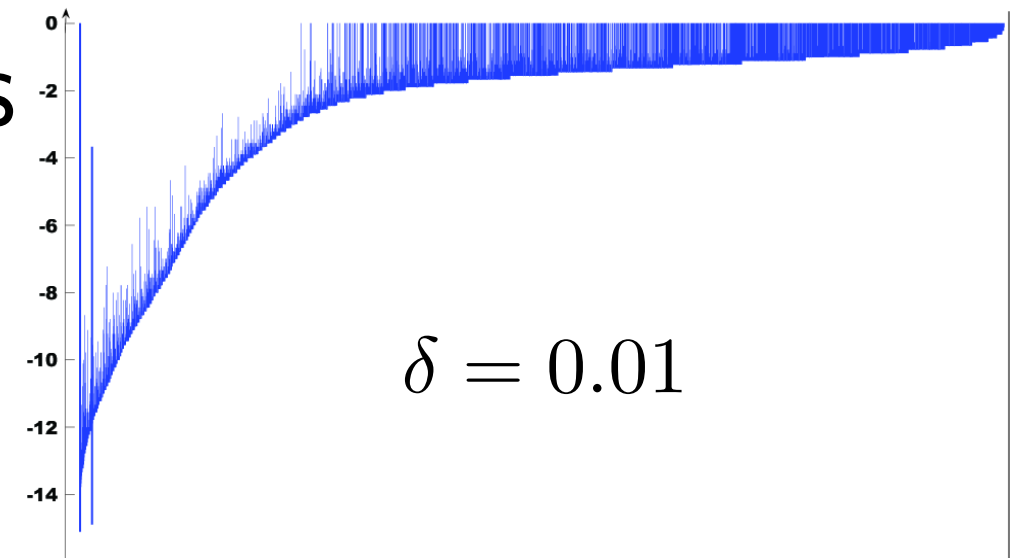
$f = \# \{ \text{data pts in fixed-radius ball} \}$



Some Results

Input: $\mathbb{X} = [0, 1]^2$; $|L| = 100,000$;

$f = \# \{ \text{data pts in fixed-radius ball} \}$



Take-Home Message(s)

- Extension of the persistence paradigm to non-triangulated or non-triangulable spaces.
- New applications of the persistence-based approach: clustering, shape segmentation, ... any problem cast into the one of finding significant peaks in some scalar field over a sampled space.
- Extensions:
 - robustness to noise,
 - partially-sampled spaces,
 - time-varying scalar fields,
 - non-manifold Alexandrov spaces.

