ATMCS 2010, Münster, June 25

Stability of Persistence Diagrams
Application to Scalar Fields Analysis

Steve Y. Oudot — Geometrica group, INRIA Saclay – Île-de-France

Some References

[1] Chazal, Cohen-Steiner, Glisse, Guibas, O., *Proximity of persistence modules and their diagrams*, Proc. SoCG 2009.

[2] Chazal, Guibas, O., Skraba, *Analysis of scalar fields over point cloud data*, Proc. SODA 2009.

[3] Chazal, Guibas, O., Skraba, "Persistence-Based Clustering in Riemannian Manifolds", INRIA Research report 6968, June 2009.

Some References

[1] Chazal, Cohen-Steiner, Glisse, Guibas, O., *Proximity of persistence modules and their diagrams*, Proc. SoCG 2009.

[2] Chazal, Guibas, O., Skraba, *Analysis of scalar fields over point cloud data*, Proc. SODA 2009.

[3] Chazal, Guibas, O., Skraba, "Persistence-Based Clustering in Riemannian Manifolds", INRIA Research report 6968, June 2009.

Logical order: 1 - 2 - 3

Some References

[1] Chazal, Cohen-Steiner, Glisse, Guibas, O., *Proximity of persistence modules and their diagrams*, Proc. SoCG 2009.

[2] Chazal, Guibas, O., Skraba, *Analysis of scalar fields over point cloud data*, Proc. SODA 2009.

[3] Chazal, Guibas, O., Skraba, "Persistence-Based Clustering in Riemannian Manifolds", INRIA Research report 6968, June 2009.

Historical order: 2 - 1 - 2 - 3

Some References

[1] Chazal, Cohen-Steiner, Glisse, Guibas, O., *Proximity of persistence modules and their diagrams*, Proc. SoCG 2009.

[2] Chazal, Guibas, O., Skraba, *Analysis of scalar fields over point cloud data*, Proc. SODA 2009.

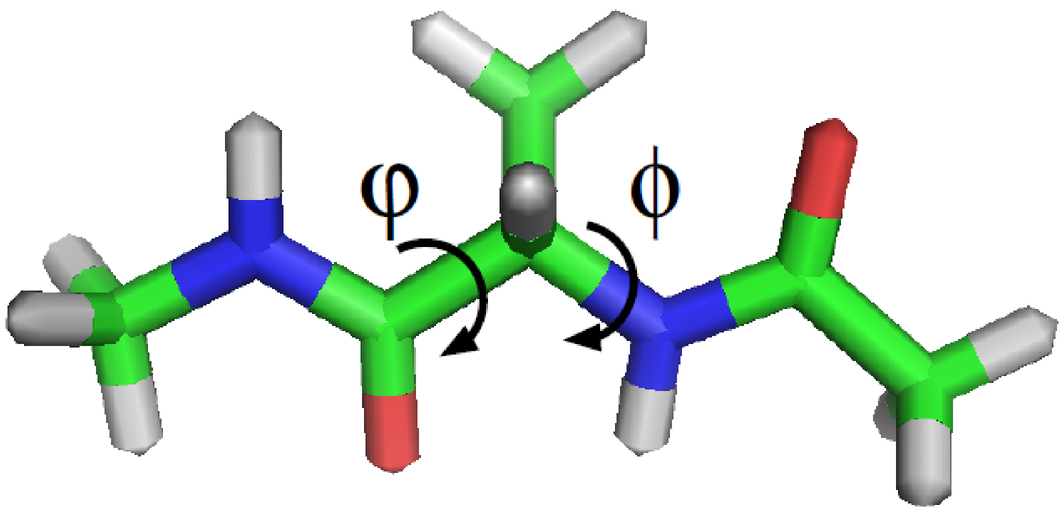
[3] Chazal, Guibas, O., Skraba, "Persistence-Based Clustering in Riemannian Manifolds", INRIA Research report 6968, June 2009.

Today's order: 3 - 2 - (1) - 2 - 3

The Alanine-Dipeptide Dataset

The alanine-dipeptide:

- 7 heavy atoms (C, N) and 13 light atoms (O, H)
- only the 7 heavy atoms play a role in conformational dynamics
  $\longrightarrow$ each conformation is represented as a point in $\mathbb{R}^{21}$
- conformation space endowed with Root Mean-Squared Deviation (RMSD)

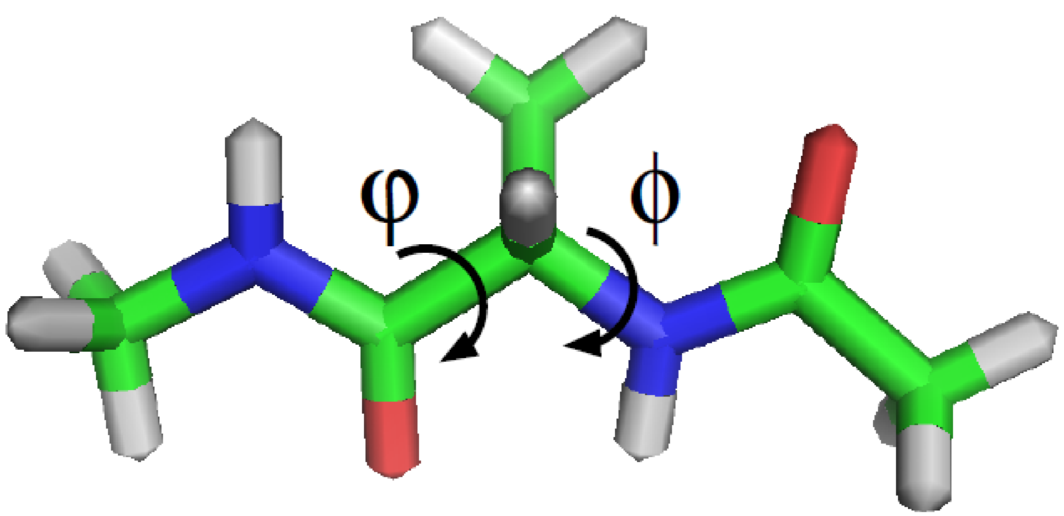


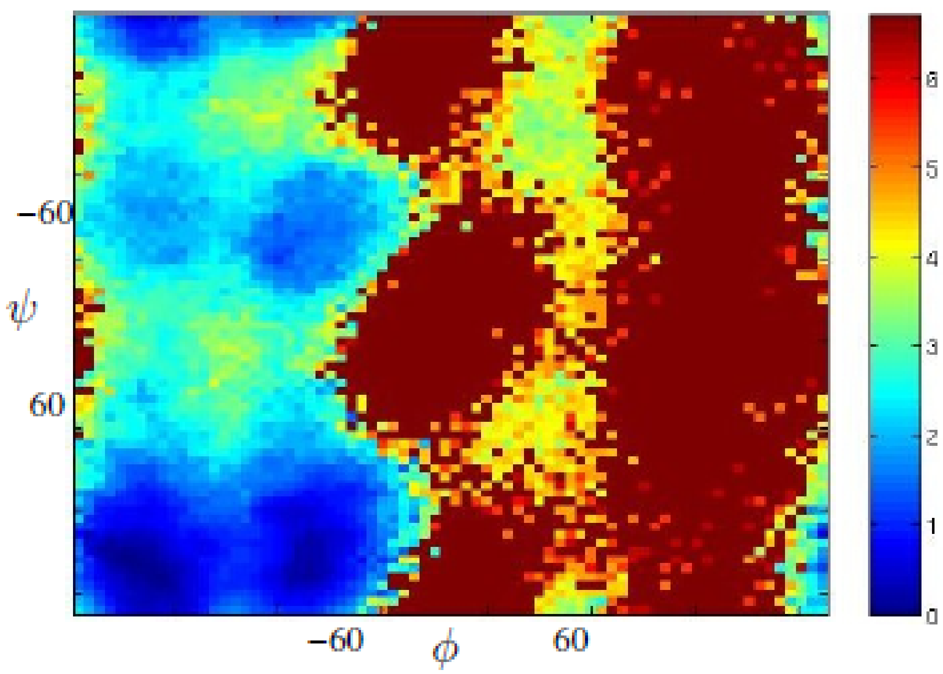
(from [Chodera et al. 2007], courtesy of the *Folding@Home* project)

The Alanine-Dipeptide Dataset

The alanine-dipeptide:

- 7 heavy atoms (C, N) and 13 light atoms (O, H)
- only the 7 heavy atoms play a role in conformational dynamics
  $\rightarrow$ each conformation is represented as a point in $\mathbb{R}^{21}$
- conformation space endowed with Root Mean-Squared Deviation (RMSD)
- only two relevant DOFs (angles ϕ, ψ)
- underlying free energy landscape exhibits 6 major minima (*metastable states*)



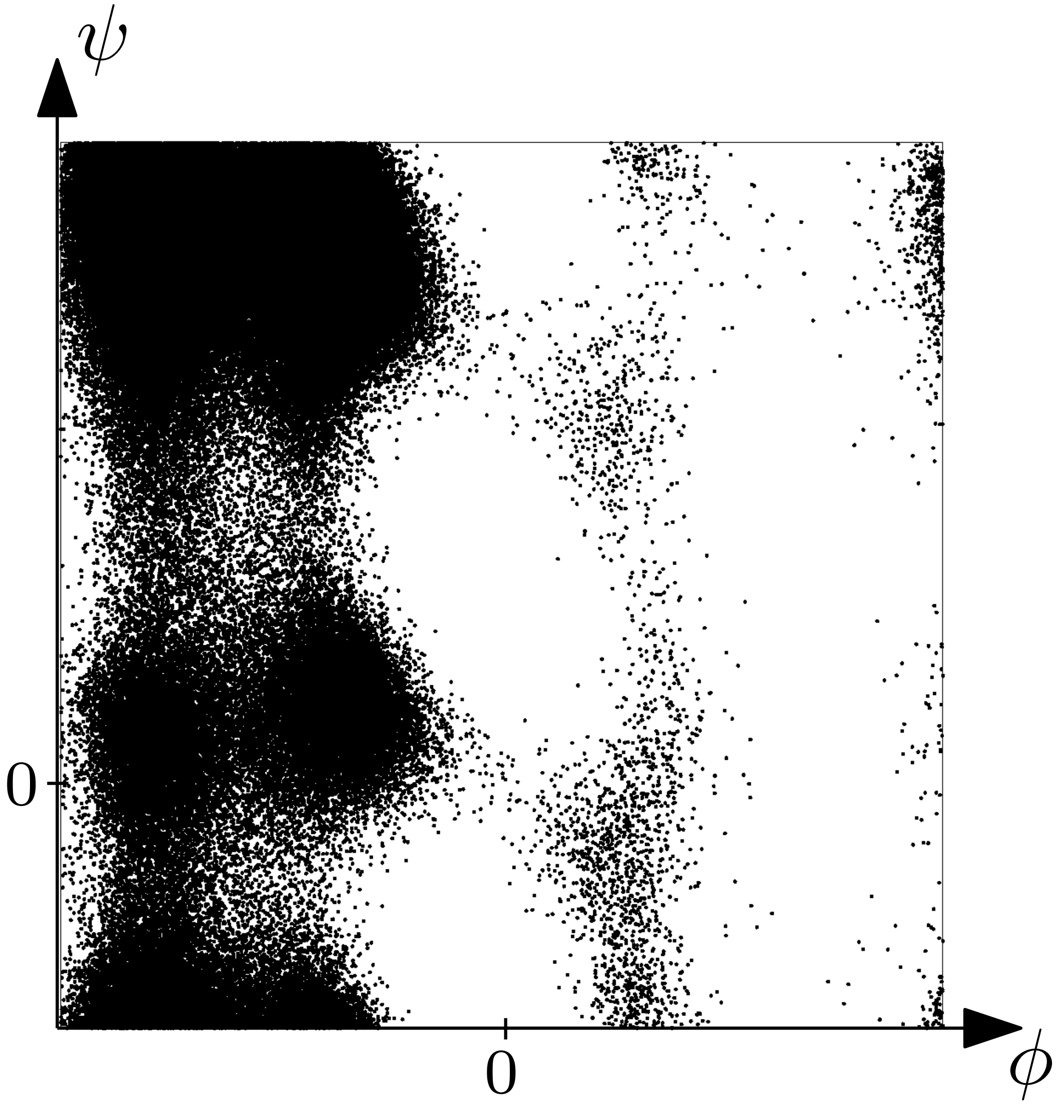


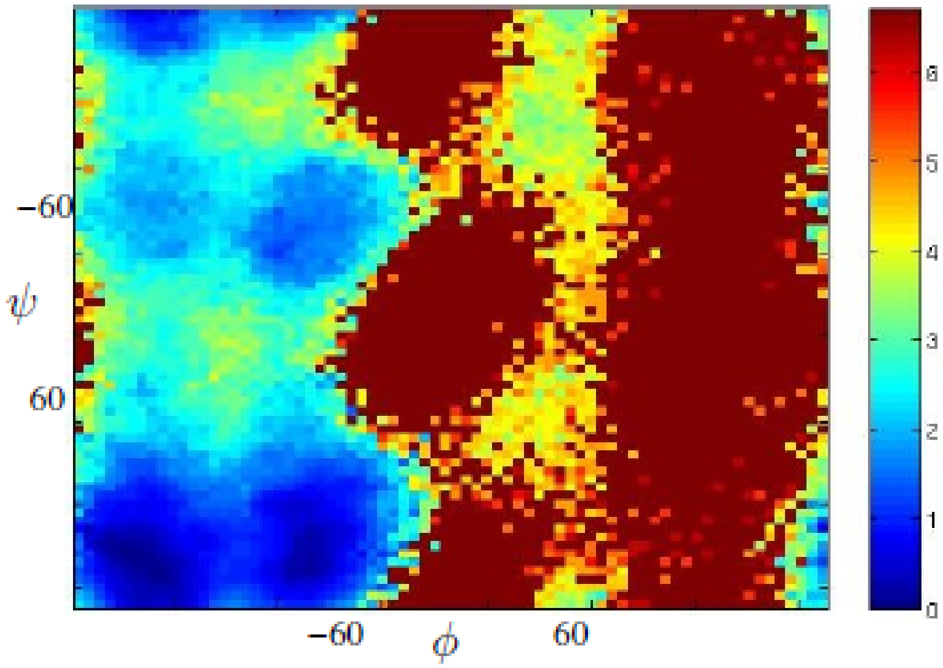
(from [Chodera et al. 2007], courtesy of the *Folding@Home* project)

The Alanine-Dipeptide Dataset

Input:

- 192,000 points in $\mathbb{R}^{21}$ obtained from simulations at picosecond scale
- RMSD used as distance between conformations

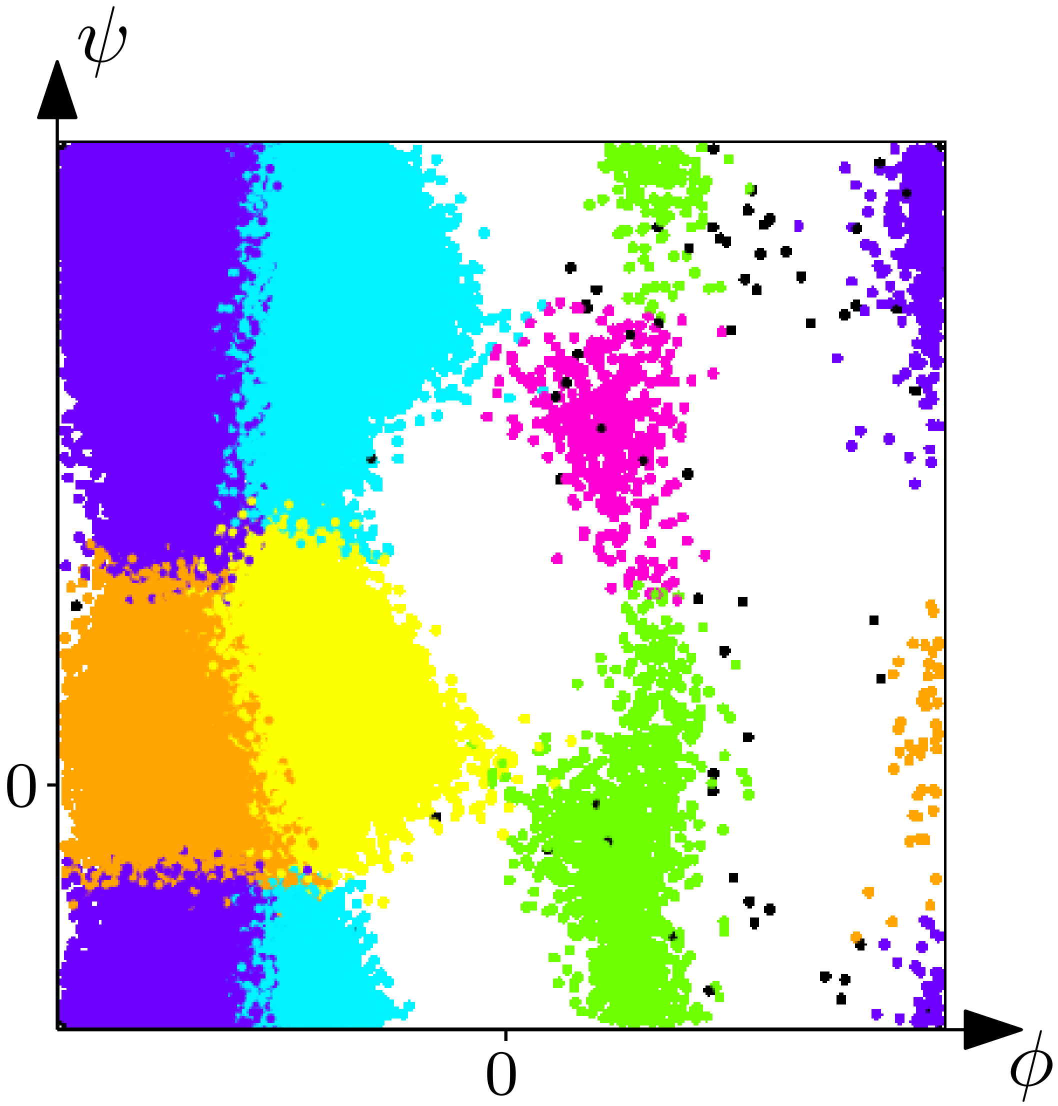




The Alanine-Dipeptide Dataset

Input:

- 192,000 points in $\mathbb{R}^{21}$ obtained from simulations at picosecond scale
- RMSD used as distance between conformations



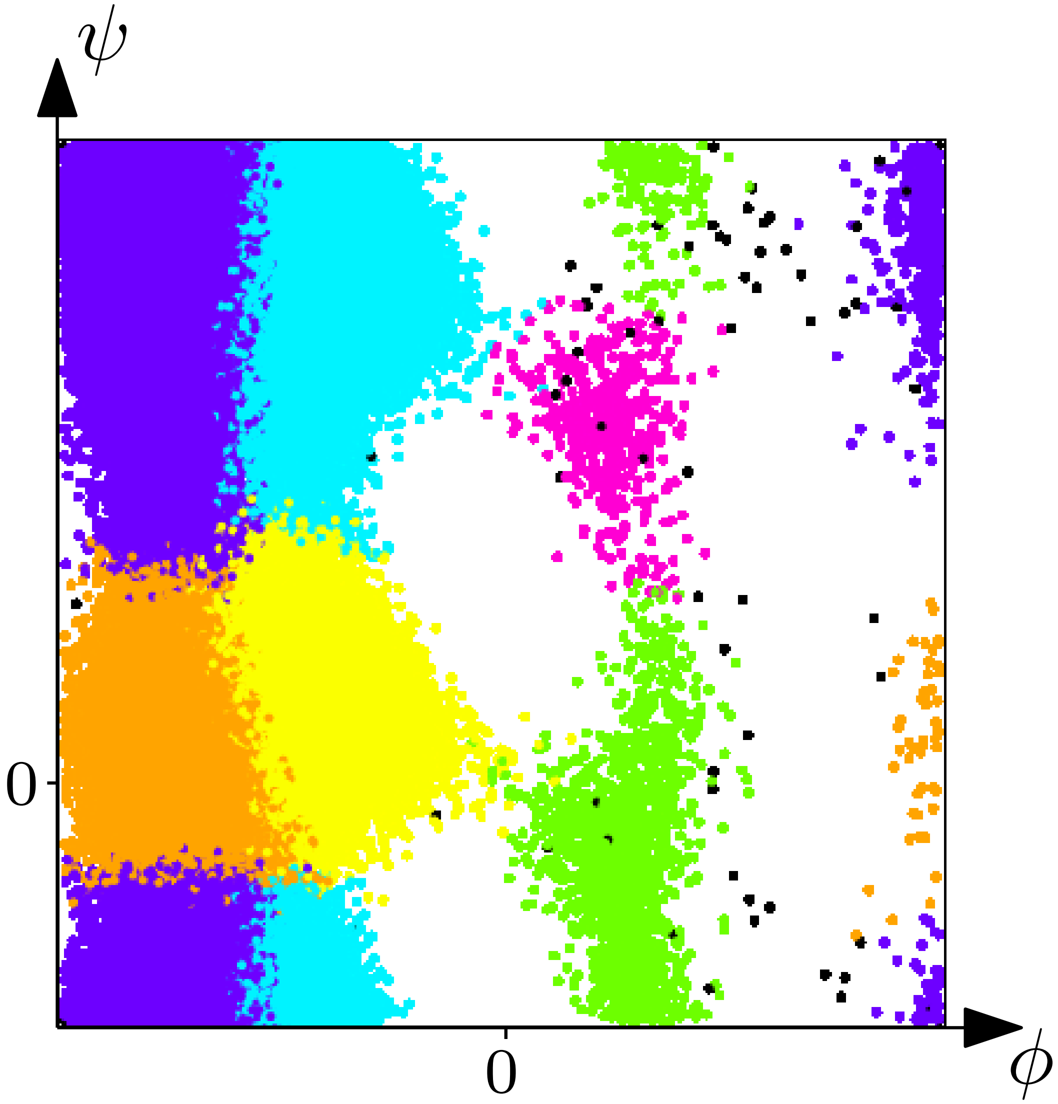
Goal:

- cluster data into *metastable states*
- states gathered into Markovian model
- enables simulations at millisecond scale

The Alanine-Dipeptide Dataset

Input:

- 192,000 points in $\mathbb{R}^{21}$ obtained from simulations at picosecond scale
- RMSD used as distance between conformations



Goal:

- cluster data into *metastable states*
- states gathered into Markovian model
- enables simulations at millisecond scale

Challenges:

- low-dim. data embedded in higher dim.
- non-trivial topology
- ambient space is non-Euclidean
- prominences of density peaks differ a lot

The Alanine-Dipeptide Dataset

Input:

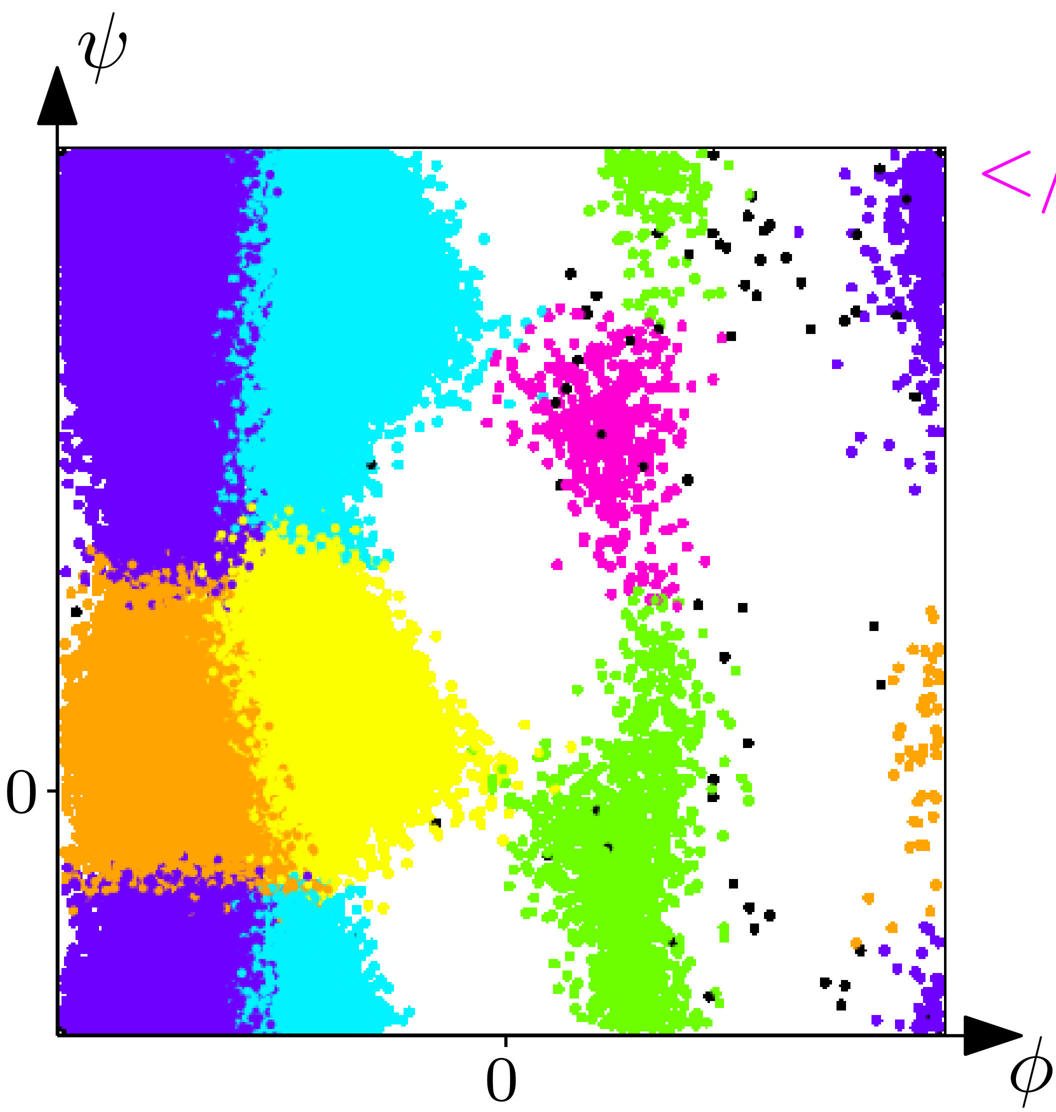
- 192,000 points in $\mathbb{R}^{21}$ obtained from simulations at picosecond scale
- RMSD used as distance between conformations

<advertising>

previous attempts took $\geq$ week of tweaking

our approach worked out of the box (10 minutes)

</advertising>

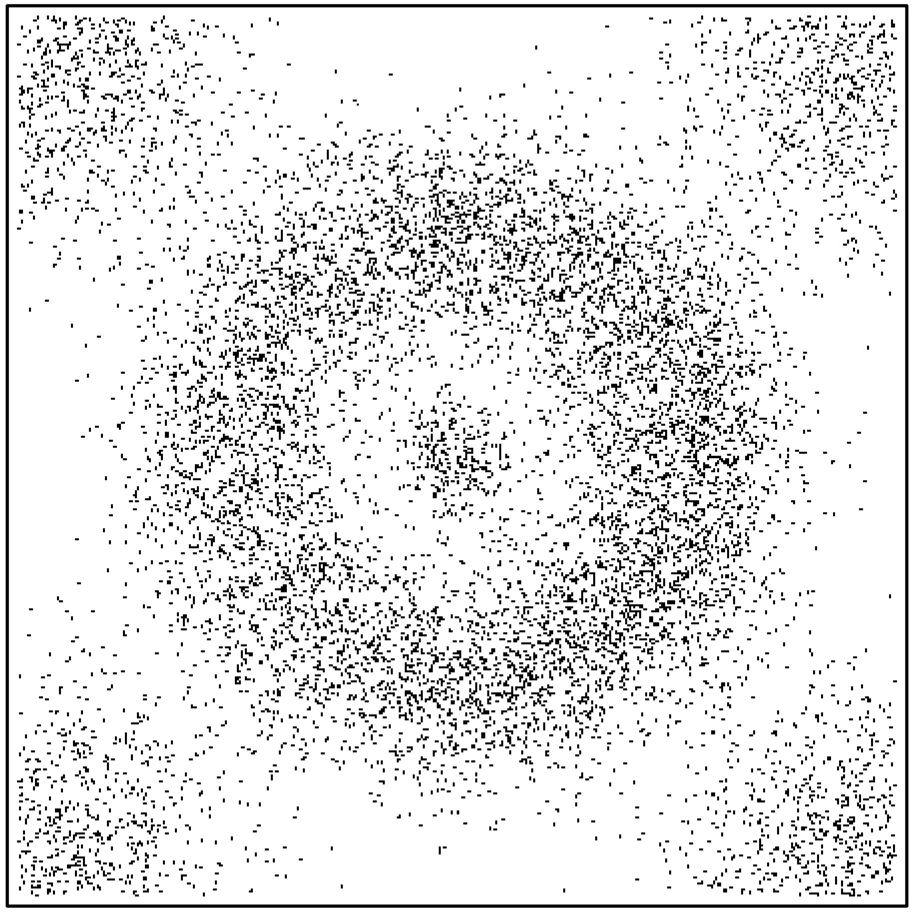


Challenges:

- low-dim. data embedded in higher dim.
- non-trivial topology
- ambient space is non-Euclidean
- prominences of density peaks differ a lot

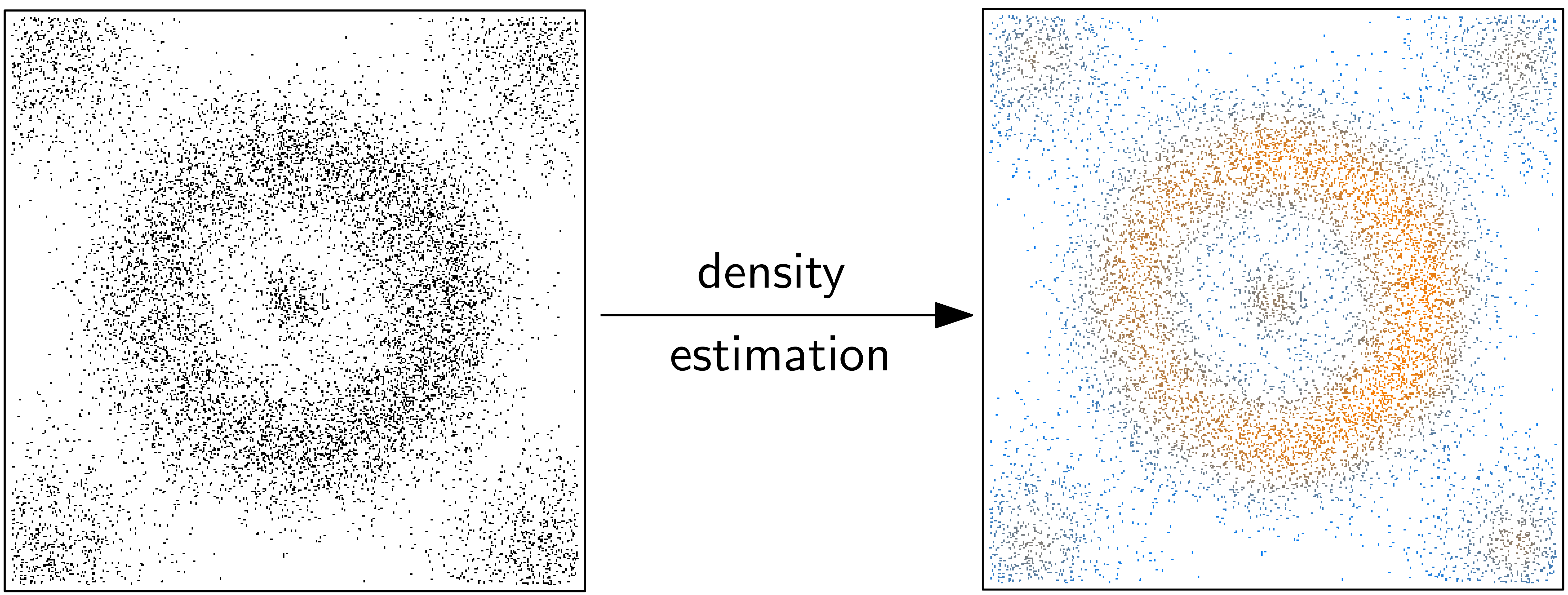
Our Formalism

- $\mathbb{X}$ is a Riemannian manifold
- points are drawn from an unknown probability distribution of density $f : \mathbb{X} \to \mathbb{R}$
- sole input: the pairwise geodesic distances between the data points



Our Formalism

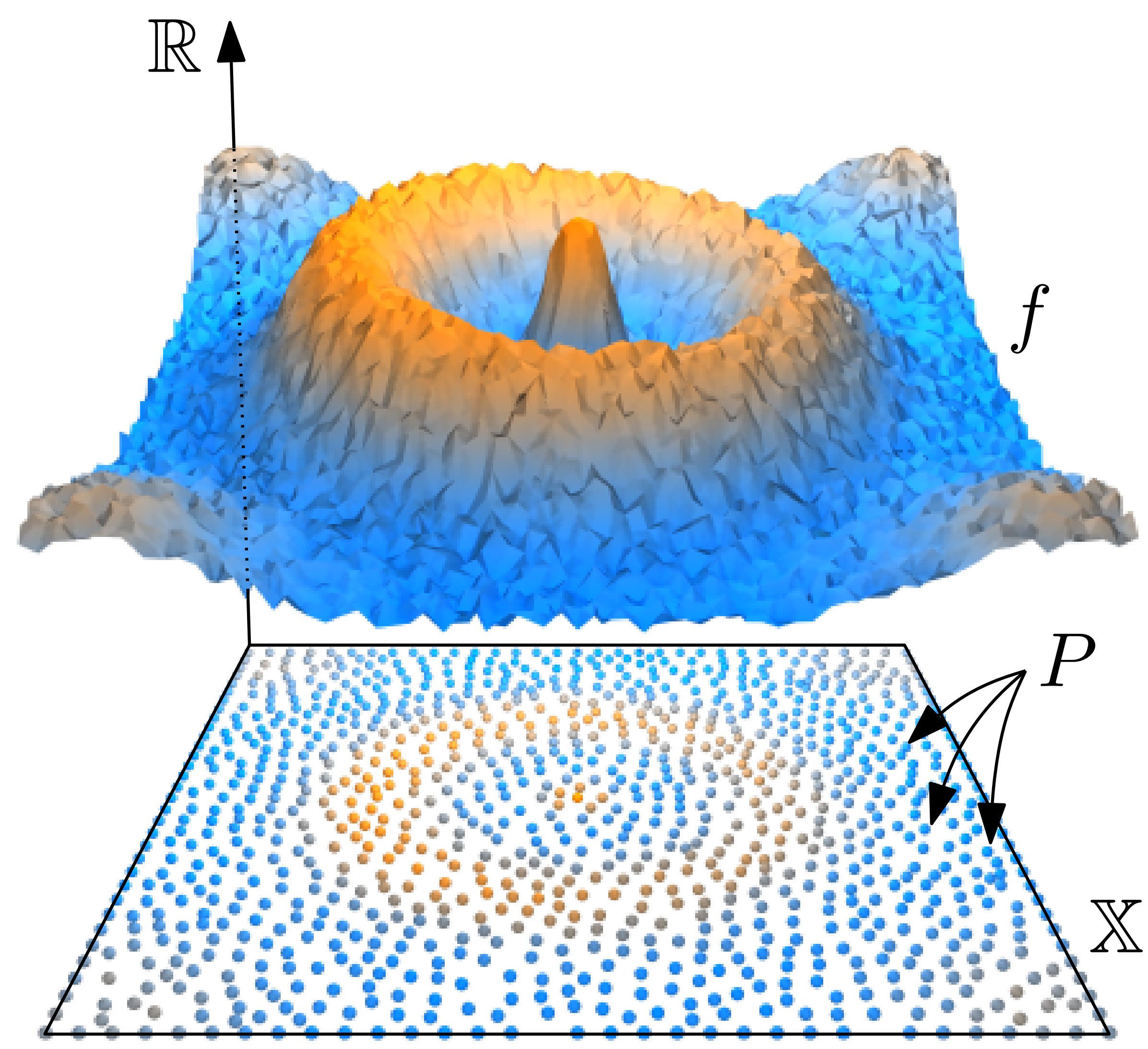
- $\mathbb{X}$ is a Riemannian manifold
- points are drawn from an unknown probability distribution of density $f : \mathbb{X} \to \mathbb{R}$
- sole input: the pairwise geodesic distances between the data points
- approximate f through some density estimator $\hat{f}$
- cluster data points according to basins of attraction of *prominent* peaks of $\hat{f}$



Link with Scalar Fields Analysis

Setting: $\mathbb{X}$ a metric space, $f : \mathbb{X} \to \mathbb{R}$ a *regular* function
P a finite point cloud in $\mathbb{X}$

Input: pairwise distances between points of P, values of f at P



Link with Scalar Fields Analysis

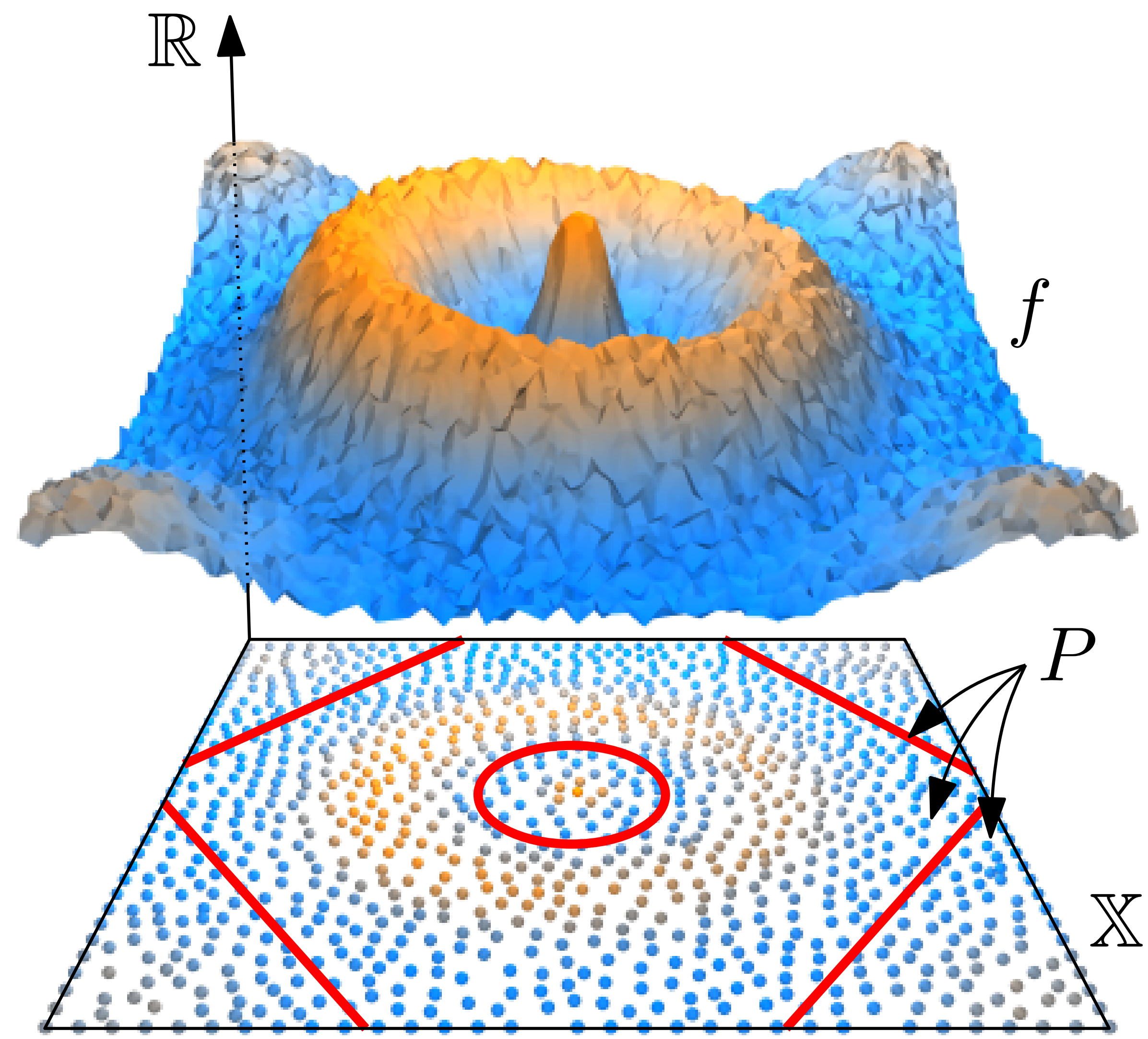
Setting: $\mathbb{X}$ a metric space, $f : \mathbb{X} \to \mathbb{R}$ a *regular* function
P a finite point cloud in $\mathbb{X}$

Input: pairwise distances between points of P, values of f at P

Goal: Analyze the topological structure of f:

- persistence diagrams

- stable/unstable manifolds



Link with Scalar Fields Analysis

Setting: $\mathbb{X}$ a metric space, $f : \mathbb{X} \to \mathbb{R}$ a *regular* function
P a finite point cloud in $\mathbb{X}$

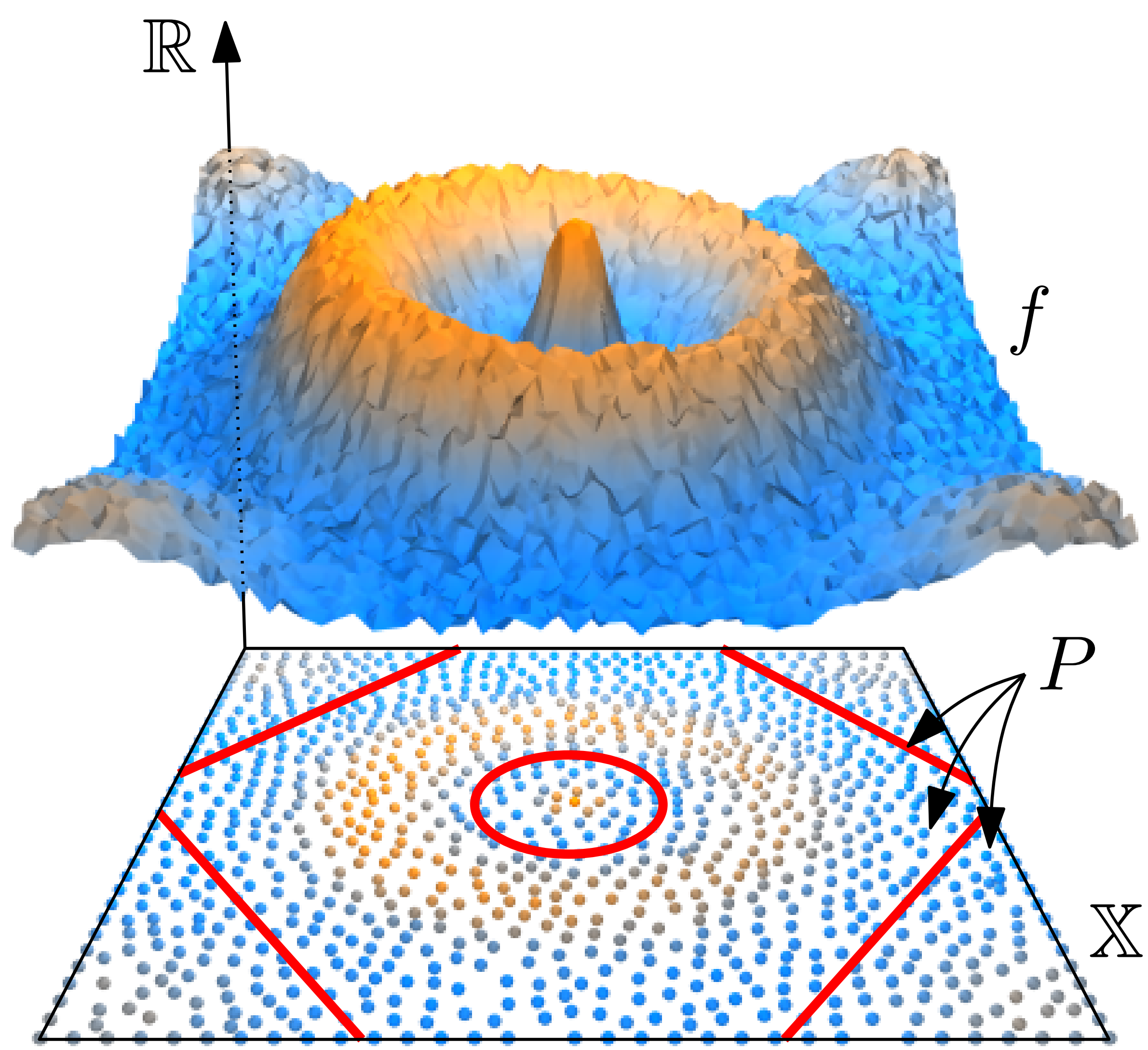
Input: pairwise distances between points of P, values of f at P

Goal: Analyze the topological structure of f:

- persistence diagrams

- stable/unstable manifolds

No triangulation of $\mathbb{X}$ is available



Unions of Balls

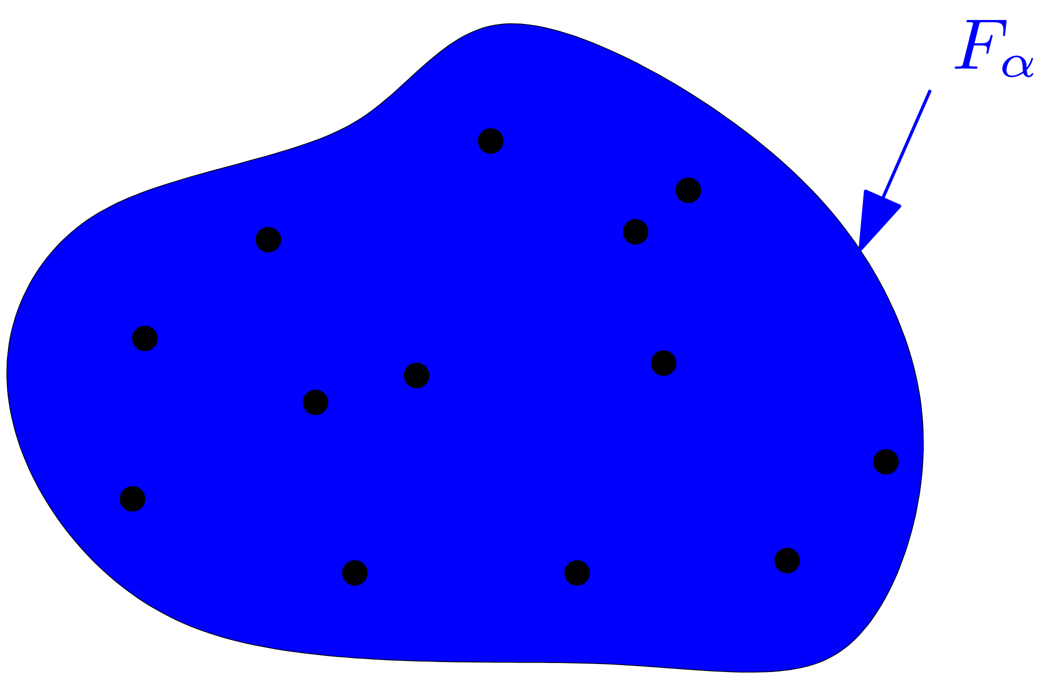
Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.

$$F_\alpha := f^{-1}((-\infty, \alpha])$$

$$P_\alpha := P \cap F_\alpha$$

$$P_\alpha^\varepsilon := \bigcup_{p \in P_\alpha} B_{\mathbb{X}}(p, \varepsilon)$$

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

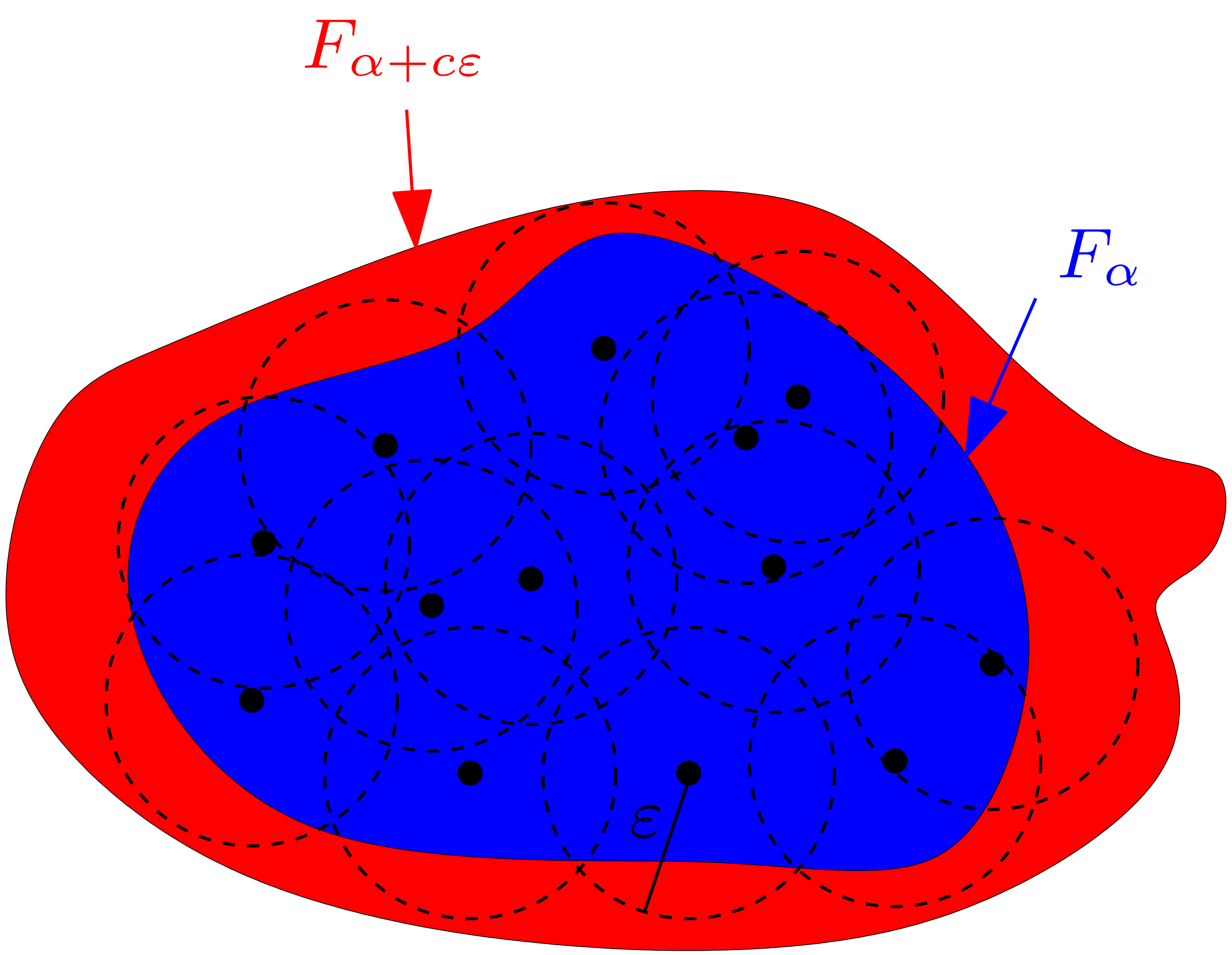


Unions of Balls

Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.

$$F_\alpha := f^{-1}((-\infty, \alpha])$$
$$P_\alpha := P \cap F_\alpha$$
$$P_\alpha^\varepsilon := \bigcup_{p \in P_\alpha} B_{\mathbb{X}}(p, \varepsilon)$$

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



Unions of Balls

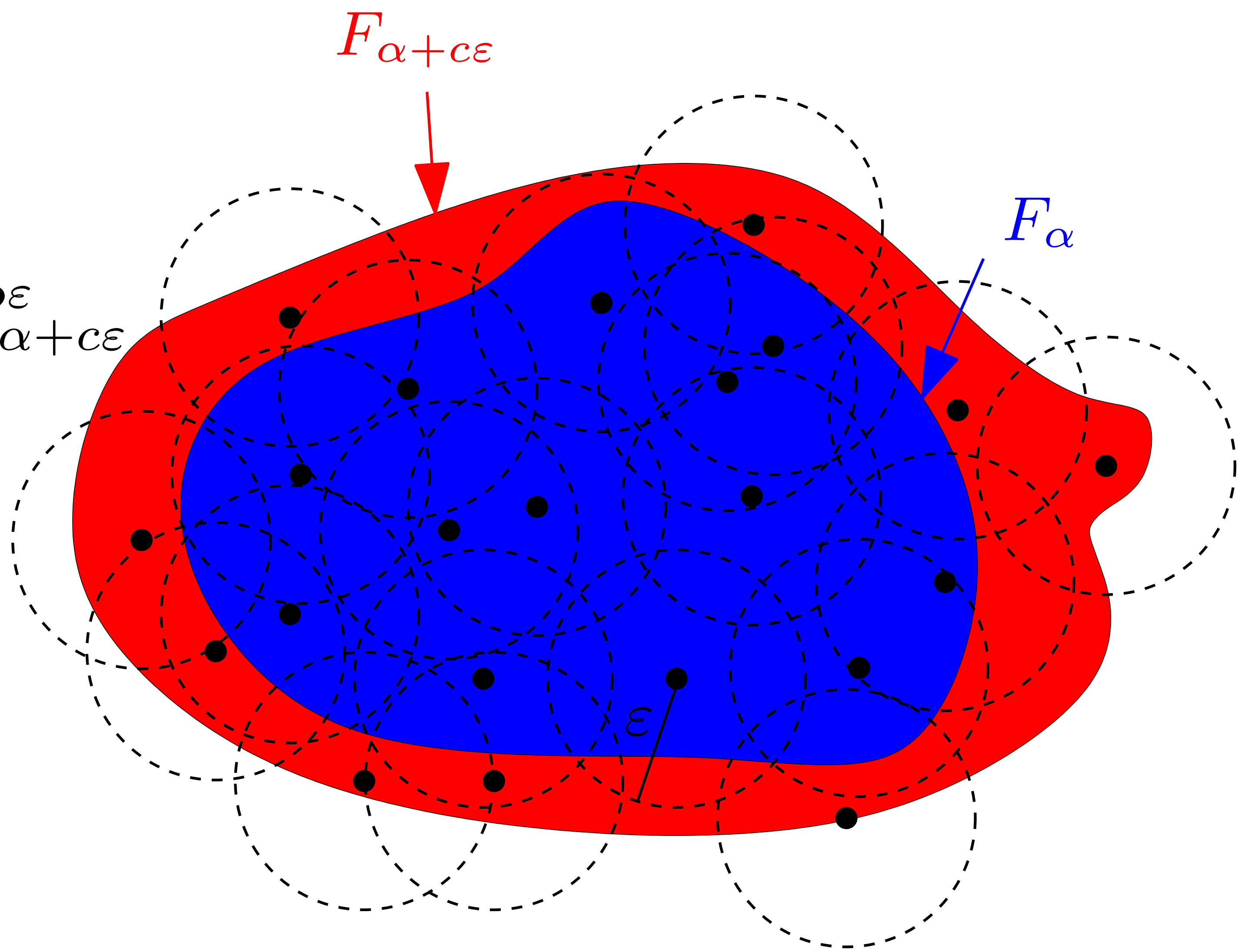
Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.

$$F_\alpha := f^{-1}((-\infty, \alpha])$$

$$P_\alpha := P \cap F_\alpha$$

$$P_\alpha^\varepsilon := \bigcup_{p \in P_\alpha} B_{\mathbb{X}}(p, \varepsilon)$$

$\forall \alpha \in \mathbb{R}$, $P_\alpha^\varepsilon \subseteq F_{\alpha + c\varepsilon}$ and $F_\alpha \subseteq P_{\alpha + c\varepsilon}^\varepsilon$



Unions of Balls

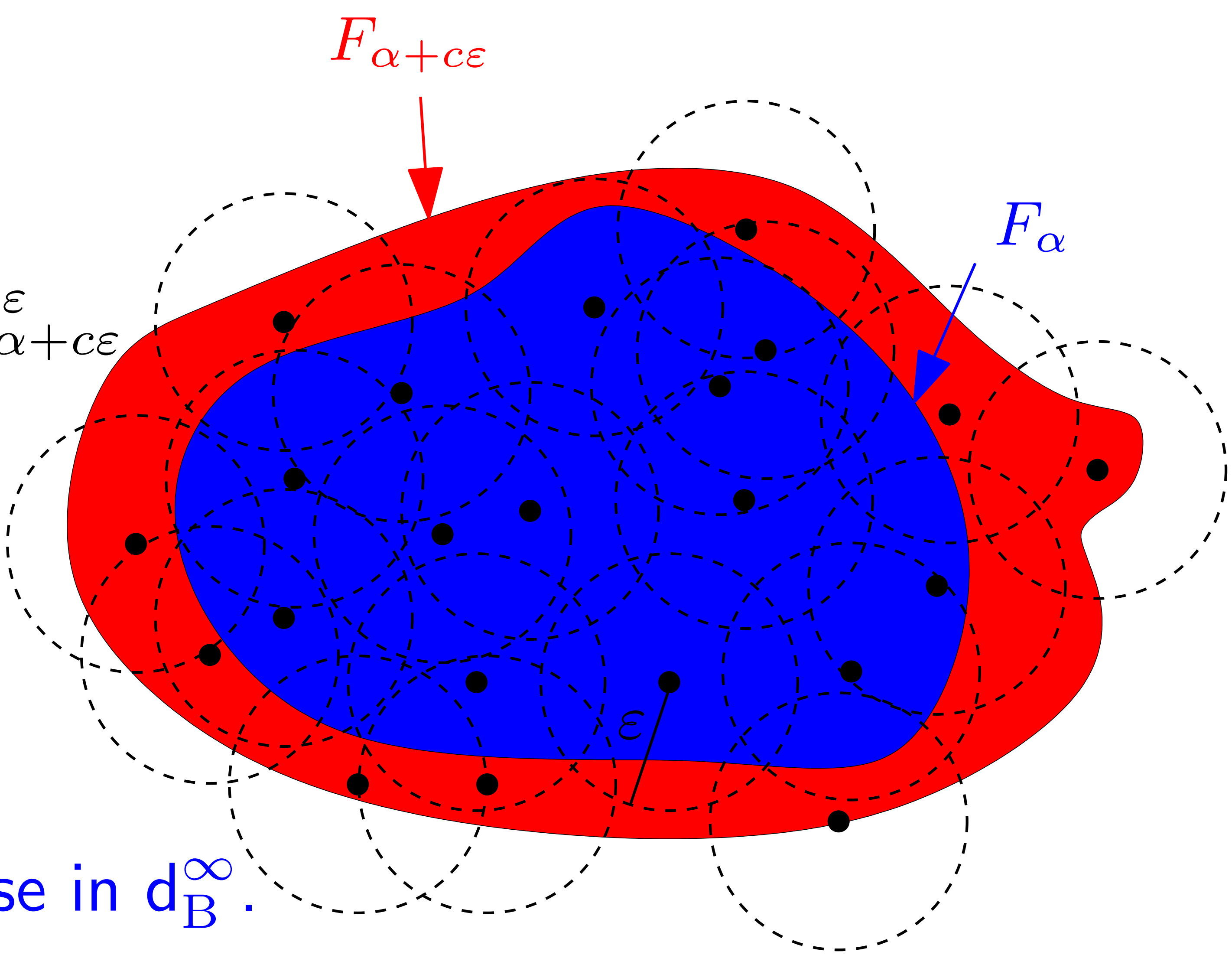
Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.

$$F_\alpha := f^{-1}((-\infty, \alpha])$$

$$P_\alpha := P \cap F_\alpha$$

$$P_\alpha^\varepsilon := \bigcup_{p \in P_\alpha} B_{\mathbb{X}}(p, \varepsilon)$$

$\forall \alpha \in \mathbb{R}$, $P_\alpha^\varepsilon \subseteq F_{\alpha+c\varepsilon}$ and $F_\alpha \subseteq P_{\alpha+c\varepsilon}^\varepsilon$



the nested families of spaces
$\{F_\alpha\}_{\alpha\in\mathbb{R}}$ and $\{P_\alpha^\varepsilon\}_{\alpha\in\mathbb{R}}$
are $c\varepsilon$-interleaved w.r.t. inclusion

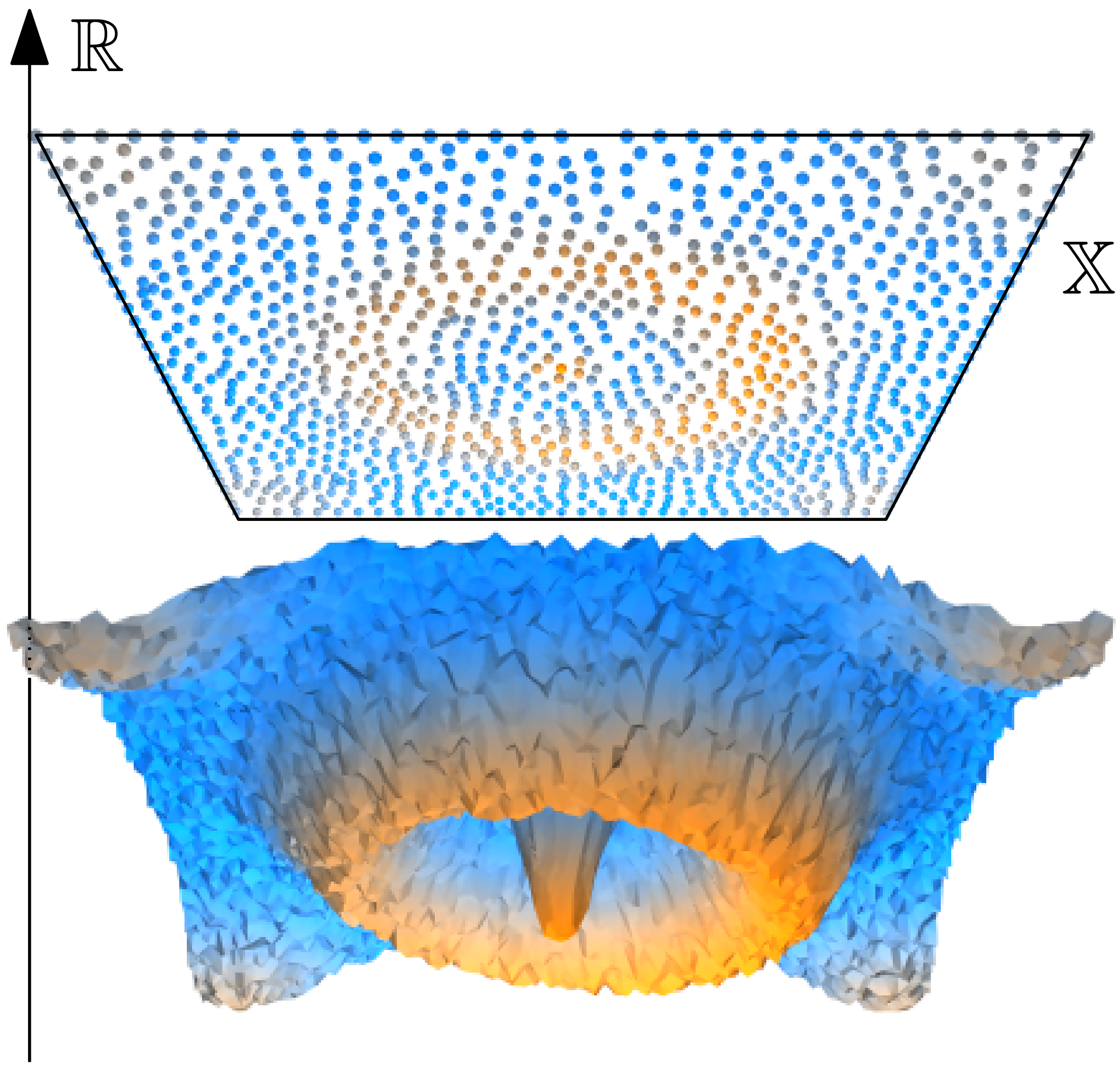
$\Downarrow$

their persistence diagrams are $c\varepsilon$-close in $\mathrm{d}_{\mathrm{B}}^{\infty}$.

[Cohen-Steiner, Edelsbrunner, Harer '05] [Chazal, Cohen-Steiner, Glisse, Guibas, O. '09]

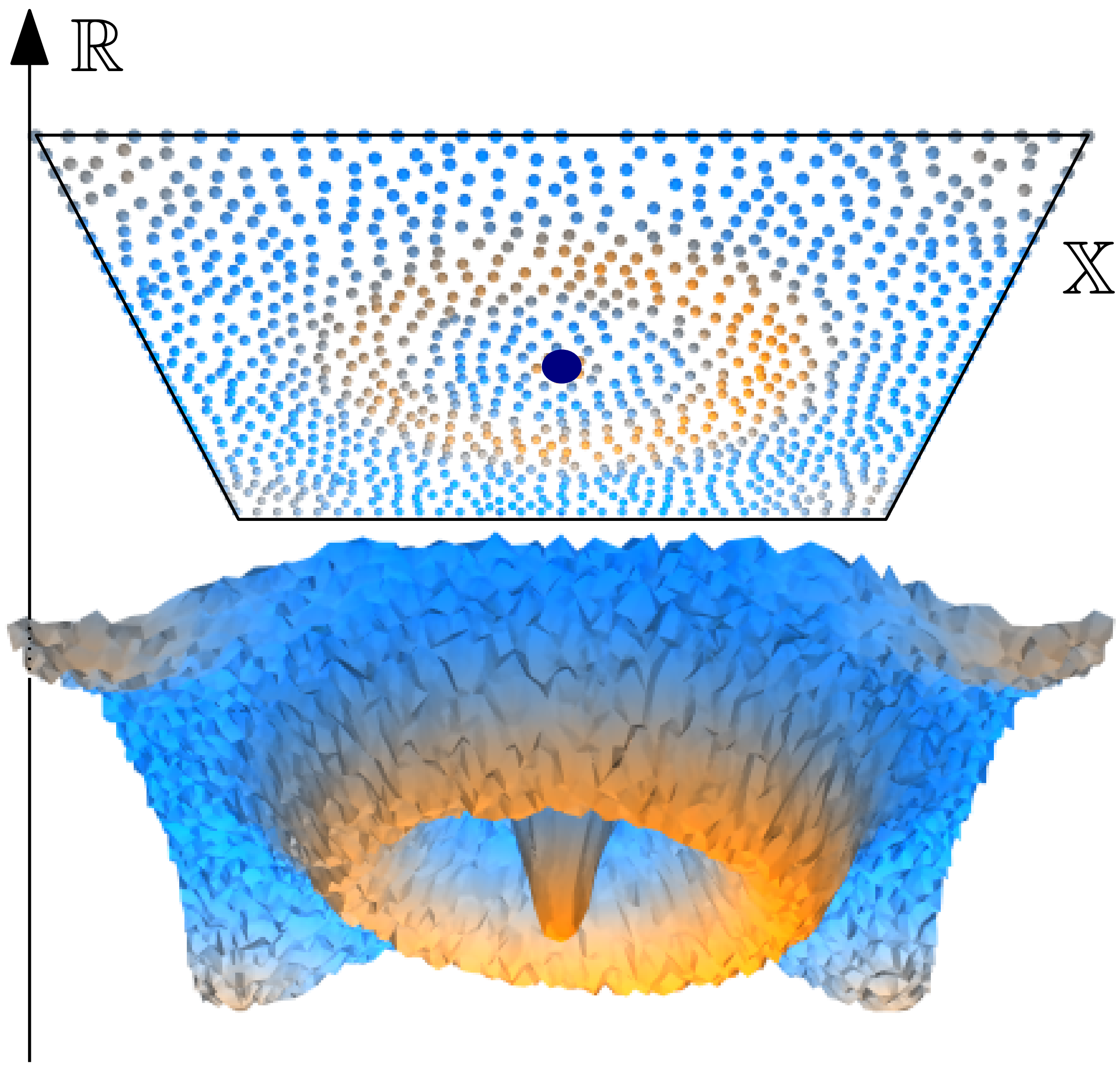
Unions of Balls

Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.



Unions of Balls

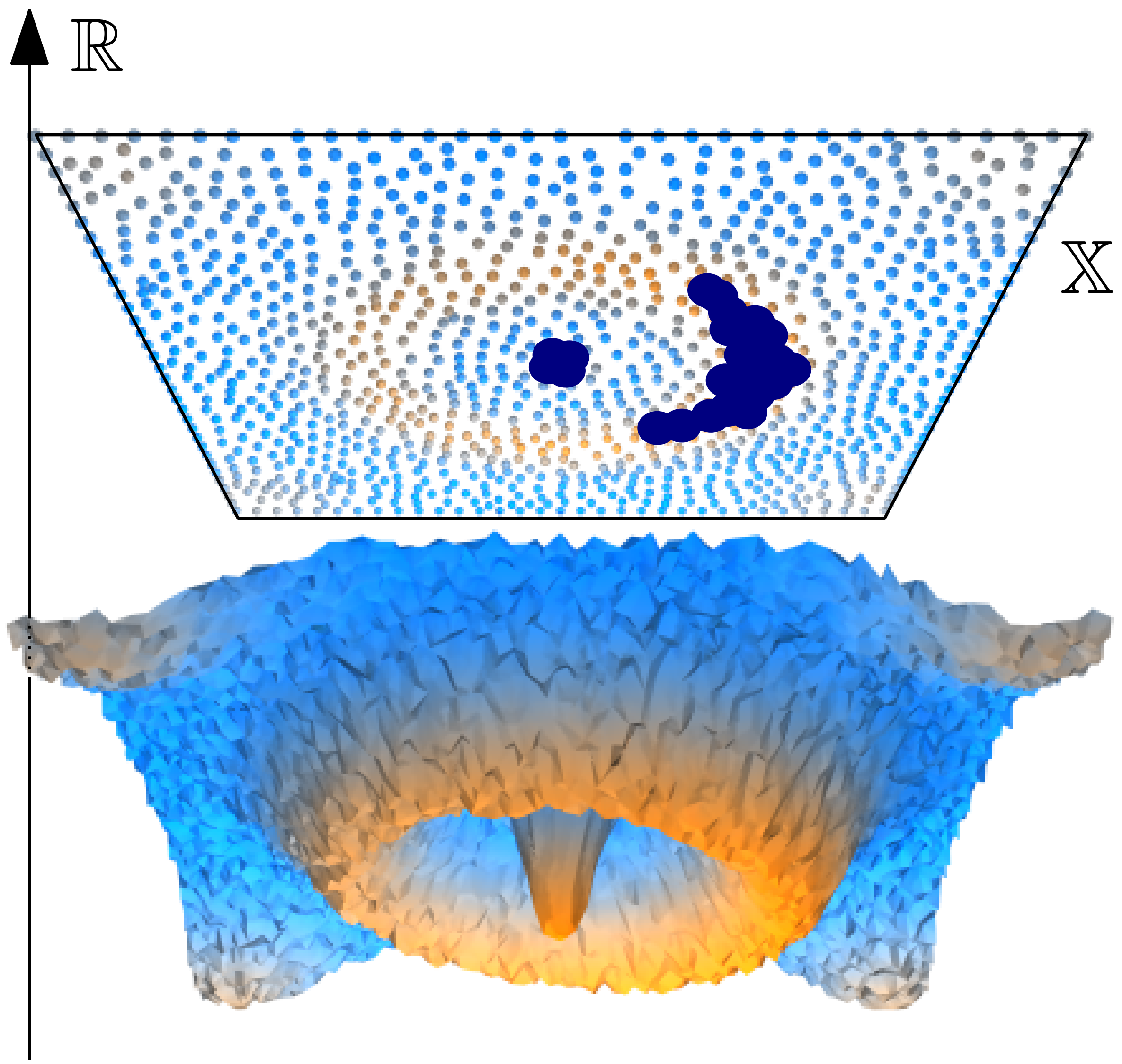
Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.



Unions of Balls

Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,

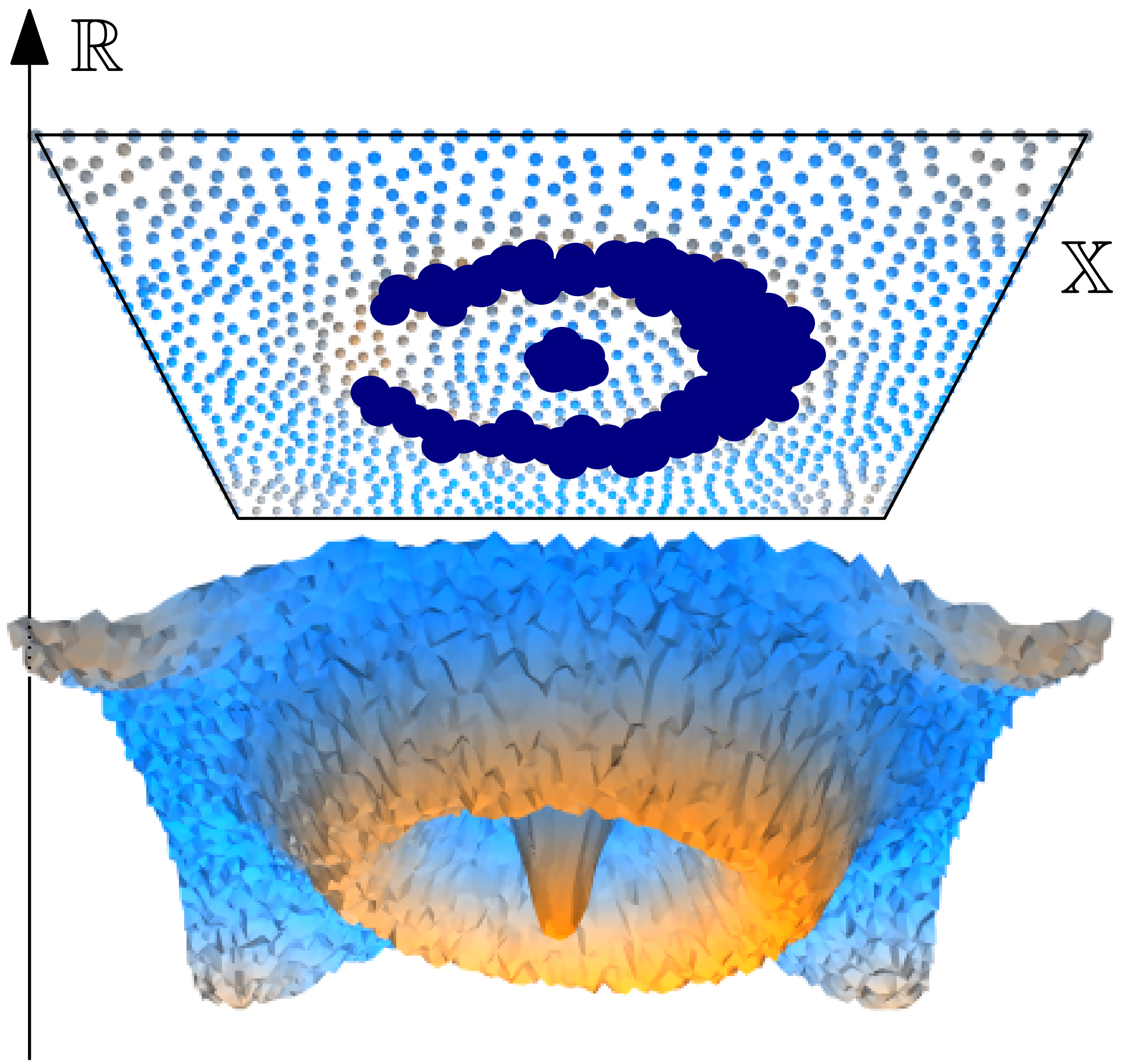
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.



Unions of Balls

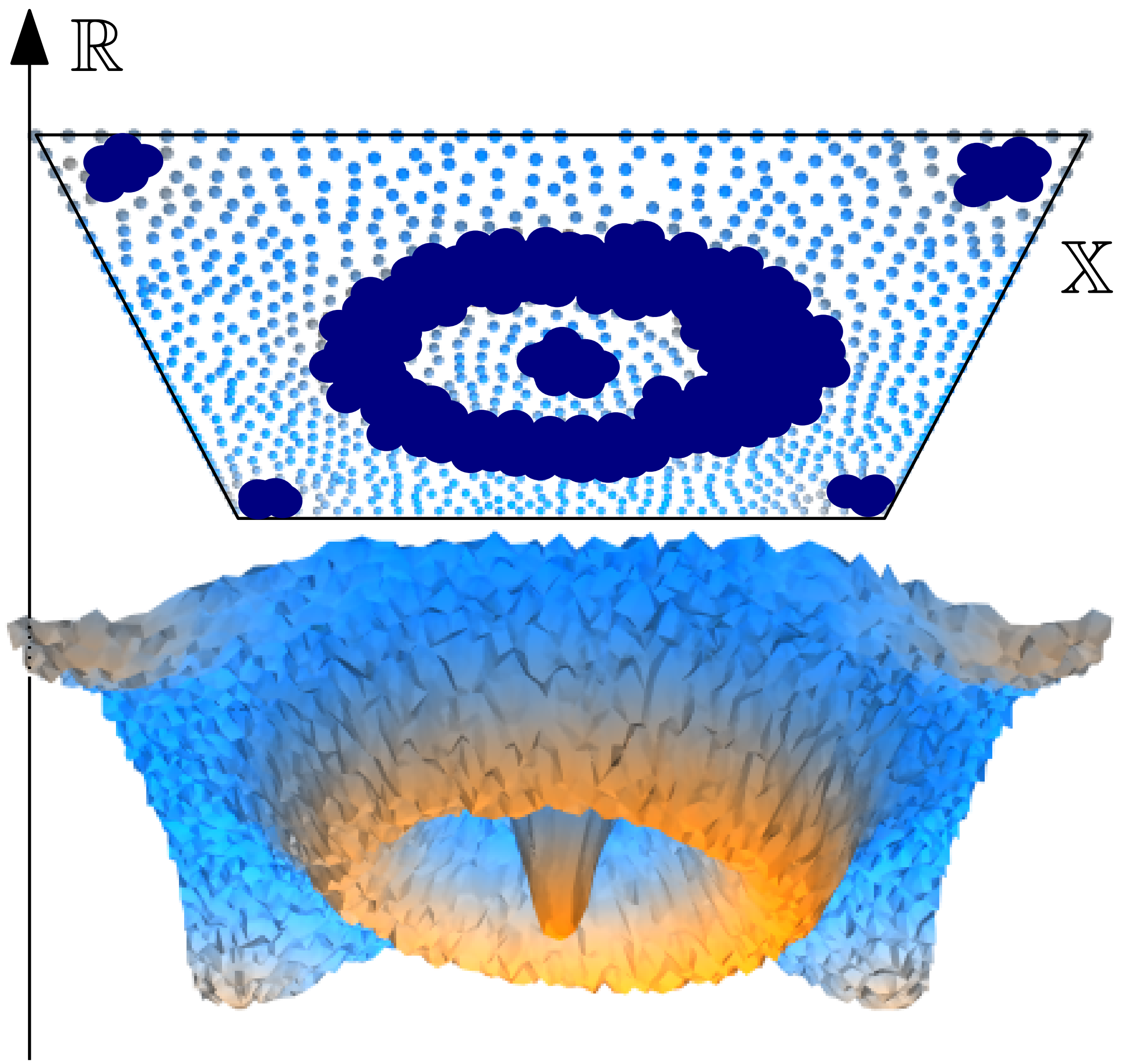
Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,

P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.



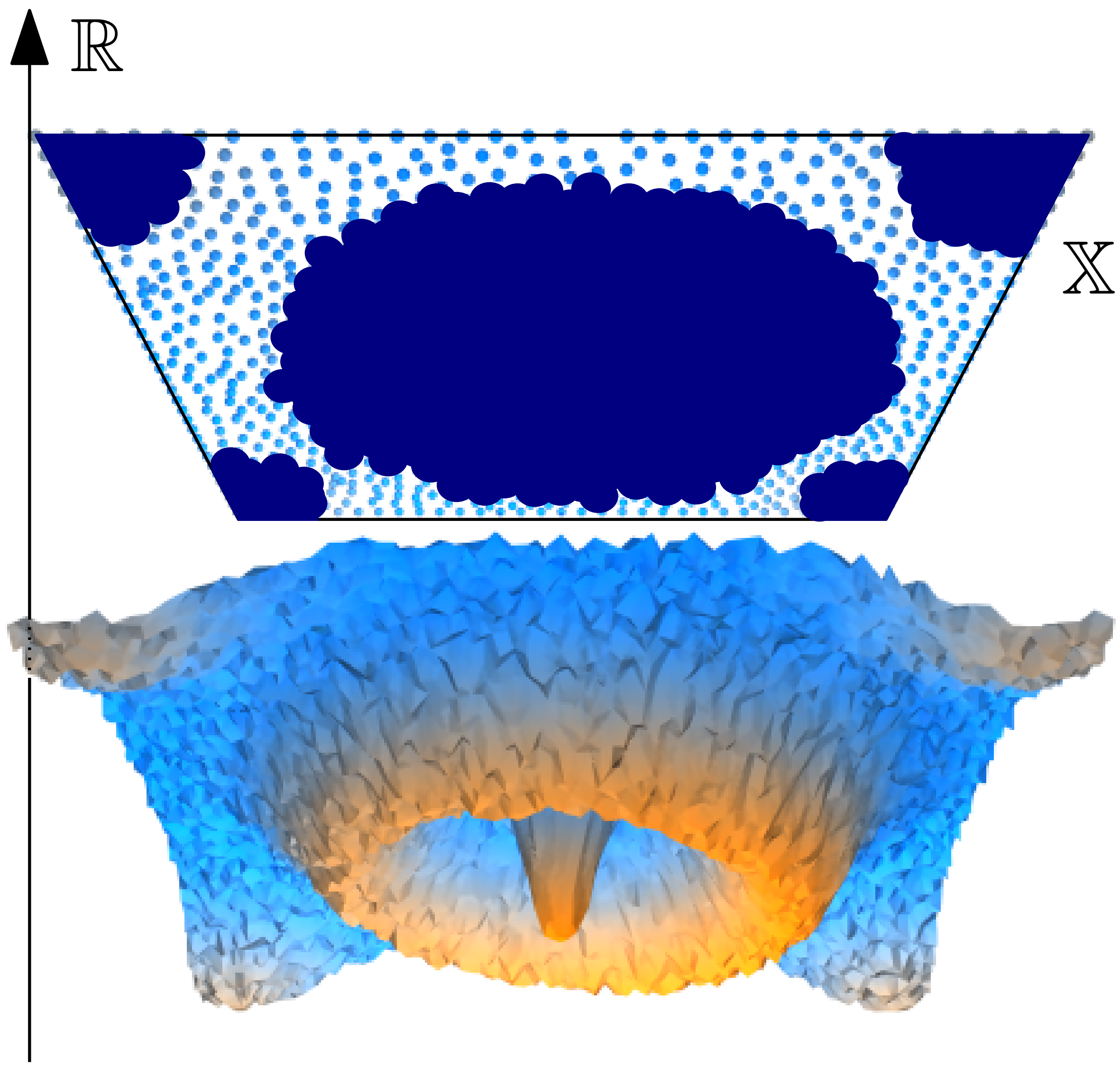
Unions of Balls

Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.



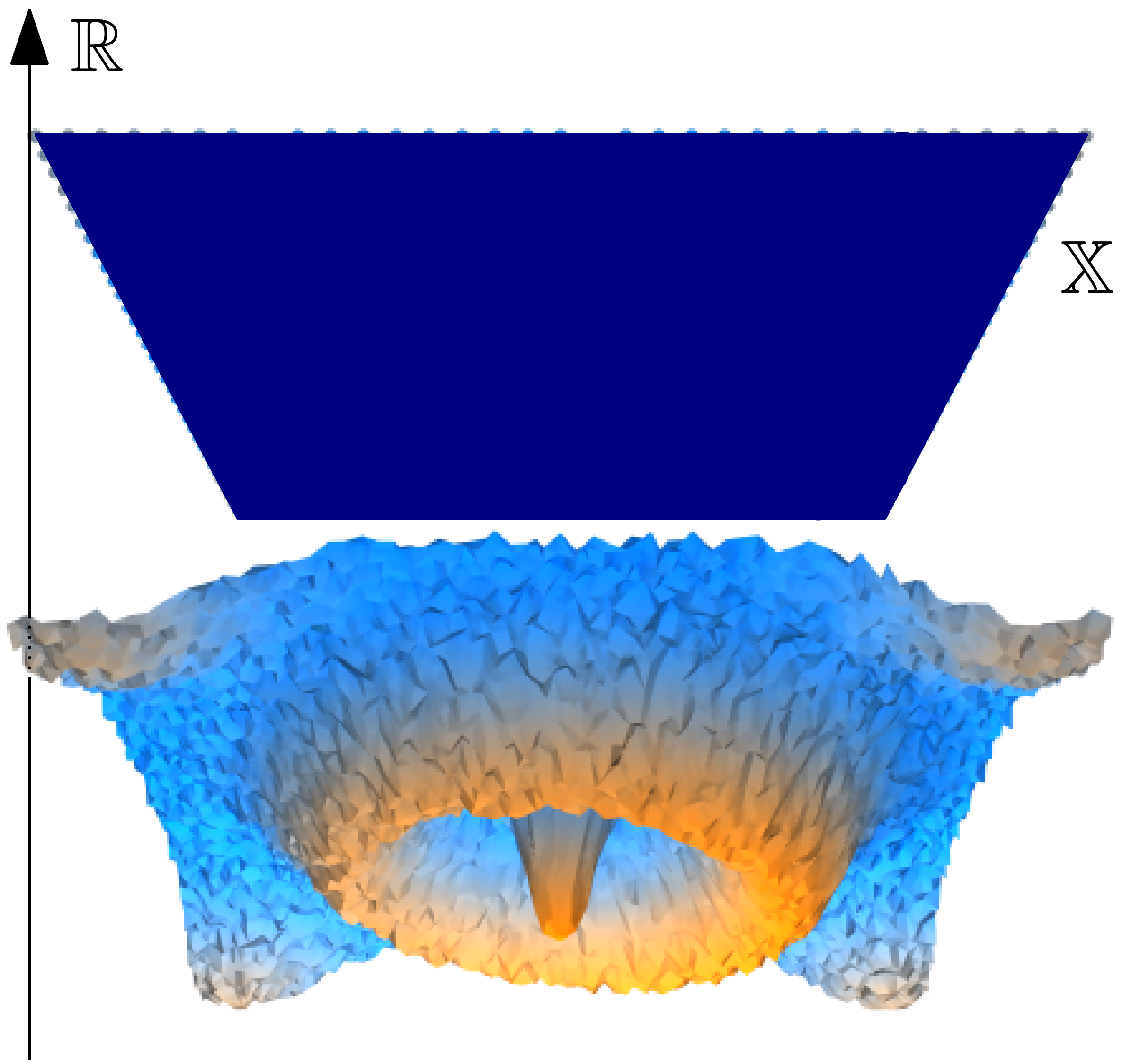
Unions of Balls

Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.



Unions of Balls

Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.



Unions of Balls

Assume $\mathbb{X}$ is a metric space, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon > 0$.

Algorithm: choose parameter $\delta \geq 0$

1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \cdots \leq f(p_n)$,
2. For $i = 1, \cdots, n$, build the union of balls $\{p_1, \cdots, p_i\}^\delta$,
3. Compute the persistent homology of the nested family of spaces:

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

Unions of Balls

Riemannian manifold

Assume $\mathbb{X}$ is a ~~metric space~~, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,

P is an ε-sample of $\mathbb{X}$ for some (unknown) $\cancel{\varepsilon > 0}$.

$0 < \varepsilon < \varrho_c(\mathbb{X})$

Algorithm: choose parameter $\delta \geq 0$

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

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

Unions of Balls

Riemannian manifold

Assume $\mathbb{X}$ is a ~~metric space~~, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,

P is an ε-sample of $\mathbb{X}$ for some (unknown) $\cancel{\varepsilon > 0}$.

$0 < \varepsilon < \varrho_c(\mathbb{X})$

Algorithm: choose parameter $\delta \geq 0$

1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \cdots \leq f(p_n)$,
2. For $i = 1, \cdots, n$, build the union of balls $\{p_1, \cdots, p_i\}^\delta$,

   or equivalently its nerve $\mathcal{N}^\delta(p_1, \cdots, p_i)$,
3. Compute the persistent homology of the nested family of spaces:

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

$\to$ if $\varepsilon \leq \delta < \varrho_c(\mathbb{X})$, we obtain a $c\delta$-approximation of the diagram of f in $\mathrm{d}_\mathrm{B}^\infty$

Unions of Balls

Assume $\mathbb{X}$ is a ~~metric space~~ Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,

P is an ε-sample of $\mathbb{X}$ for some (unknown) $\varepsilon < \varrho_c(\mathbb{X})$

Algorithm: choose parameter $\delta \geq 0$

1. Sort the data points such that $f(p_1) \leq \cdots \leq f(p_n)$,
2. For $i = 1, \cdots, n$, build the union of balls $\{p_1, \cdots, p_i\}^\delta$,

   or equivalently, compute $\mathcal{N}^\delta(p_1, \cdots, p_i)$,
3. Compute the persistent homology of the nested family of spaces:

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

Intractable in practice

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

$\longrightarrow$ apply sandwitching idea from [Chazal, O. 08]:

$\forall i = 1, \cdots, n$, replace $\mathcal{N}^\delta(p_1, \cdots, p_i)$ by

$$\mathcal{R}^\delta(p_1, \cdots, p_i) \subseteq \mathcal{N}^\delta(p_1, \cdots, p_i) \subseteq \mathcal{R}^{2\delta}(p_1, \cdots, p_i)$$

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Algorithm: choose parameter $\delta \geq 0$

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

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

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

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Algorithm: choose parameter $\delta \geq 0$

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

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

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

$$\boxed{\{\mathrm{im}\ H_k(\mathcal{R}^{\delta}(p_1, \cdots, p_i)) \to H_k(\mathcal{R}^{2\delta}(p_1, \cdots, p_i))\}_{1 \leq i \leq n}}$$

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Guarantees: $\forall \delta \in [2\varepsilon, \frac{1}{2}\varrho_c(\mathbb{X}))$,

$\{H_k(F_\alpha)\}_{\alpha\in\mathbb{R}}$ and $\{\mathrm{im}\ H_k(\mathcal{R}^\delta(P_\alpha)) \to H_k(\mathcal{R}^{2\delta}(P_\alpha))\}_{\alpha\in\mathbb{R}}$ are $2c\delta$-*interleaved*:

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Guarantees: $\forall \delta \in [2\varepsilon, \frac{1}{2}\varrho_c(\mathbb{X}))$,

$\{H_k(F_\alpha)\}_{\alpha\in\mathbb{R}}$ and $\{\mathrm{im}\ H_k(\mathcal{R}^\delta(P_\alpha)) \to H_k(\mathcal{R}^{2\delta}(P_\alpha))\}_{\alpha\in\mathbb{R}}$ are $2c\delta$-*interleaved*:

Letting $M_\alpha = H_k(F_\alpha)$ and $N_\alpha = \mathrm{im}\ H_k(\mathcal{R}^\delta(P_{\alpha+2c\delta})) \to H_k(\mathcal{R}^{2\delta}(P_{\alpha+2c\delta}))$,

$\exists\ \{\phi_\alpha : M_\alpha \to N_{\alpha+2c\delta}\}_{\alpha\in\mathbb{R}}$ and $\{\psi_\alpha : N_\alpha \to M_{\alpha+2c\delta}\}_{\alpha\in\mathbb{R}}$

s.t. the following diagrams commute $\forall \alpha \le \beta$:

$$\begin{tikzcd}
M_{\alpha-2c\delta} \arrow[rrr] \arrow[dr, "\phi_{\alpha-2c\delta}"] & & & M_{\beta+2c\delta} \\
& N_\alpha \arrow[r] & N_\beta \arrow[ur, "\psi_\beta"] &
\end{tikzcd}
\qquad
\begin{tikzcd}
& M_{\alpha+2c\delta} \arrow[r] & M_{\beta+2c\delta} \\
N_\alpha \arrow[ur, "\psi_\alpha"] \arrow[r] & N_\beta \arrow[ur, "\psi_\beta"] &
\end{tikzcd}$$

$$\begin{tikzcd}
& M_\alpha \arrow[r] & M_\beta \arrow[dr, "\phi_\beta"] & \\
N_{\alpha-2c\delta} \arrow[ur, "\psi_{\alpha-2c\delta}"] \arrow[rrr] & & & N_{\beta+2c\delta}
\end{tikzcd}
\qquad
\begin{tikzcd}
M_\alpha \arrow[r] \arrow[dr, "\phi_\alpha"] & M_\beta \arrow[dr, "\phi_\beta"] & \\
& N_{\alpha+2c\delta} \arrow[r] & N_{\beta+2c\delta}
\end{tikzcd}$$

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Guarantees: $\forall \delta \in [2\varepsilon, \frac{1}{2}\varrho_c(\mathbb{X}))$,

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

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Guarantees: $\forall \delta \in [2\varepsilon, \frac{1}{2}\varrho_c(\mathbb{X}))$,

$$\begin{tikzcd}
\mathcal{N}^{2\delta}(p_1) \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1, p_2) \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1, p_2, \cdots, p_n) \\
\mathcal{R}^{2\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{R}^{2\delta}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{2\delta}(p_1, p_2, \cdots, p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{N}^{\delta}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta}(p_1, p_2, \cdots, p_n) \arrow[u, hook] \\
\mathcal{R}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{R}^{\delta}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{\delta}(p_1, p_2, \cdots, p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta/2}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{N}^{\delta/2}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta/2}(p_1, p_2, \cdots, p_n) \arrow[u, hook]
\end{tikzcd}$$

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Guarantees: $\forall \delta \in [2\varepsilon, \frac{1}{2}\varrho_c(\mathbb{X}))$,

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

$$\simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq$$

$$\color{blue}{F_{\alpha_0}} \subseteq \{p_1\}^{\delta/2} \subseteq \{p_1\}^{2\delta} \subseteq \color{blue}{F_{\alpha_1}} \subseteq \cdots\cdots \subseteq \color{blue}{F_{\alpha_{n-1}}} \subseteq \{p_1,\cdots,p_n\}^{\delta/2} \subseteq \{p_1,\cdots,p_n\}^{2\delta}$$

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Guarantees: $\forall \delta \in [2\varepsilon, \frac{1}{2}\varrho_c(\mathbb{X}))$,

$\psi_{\alpha_1 - 2c\delta}$

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

ϕ_{α_0} $\simeq$ $\simeq$ $\simeq$ $\simeq$ $\simeq$ $\simeq$

$$F_{\alpha_0} \subseteq \{p_1\}^{\delta/2} \subseteq \{p_1\}^{2\delta} \subseteq F_{\alpha_1} \subseteq \cdots\cdots \subseteq F_{\alpha_{n-1}} \subseteq \{p_1, \cdots, p_n\}^{\delta/2} \subseteq \{p_1, \cdots, p_n\}^{2\delta}$$

Pairs of Rips Complexes

$$
\begin{tikzcd}
 & M_\alpha \arrow[r] & M_\beta \arrow[dr] & \\
N_{\alpha-2c\delta} \arrow[ur] \arrow[rrr] & & & N_{\beta+2c\delta}
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathcal{N}^{2\delta}(p_1) \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1,p_2) \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1,p_2,\cdots,p_n) \\
\mathcal{R}^{2\delta}(p_1) \arrow[u, hook] \arrow[r, hook] & \mathcal{R}^{2\delta}(p_1,p_2) \arrow[u, hook] \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{2\delta}(p_1,p_2,\cdots,p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta}(p_1) \arrow[u, hook] \arrow[r, hook] & \mathcal{N}^{\delta}(p_1,p_2) \arrow[u, hook] \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta}(p_1,p_2,\cdots,p_n) \arrow[u, hook] \\
\mathcal{R}^{\delta}(p_1) \arrow[u, hook] \arrow[r, hook] & \mathcal{R}^{\delta}(p_1,p_2) \arrow[u, hook] \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{\delta}(p_1,p_2,\cdots,p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta/2}(p_1) \arrow[u, hook] \arrow[r, hook] & \mathcal{N}^{\delta/2}(p_1,p_2) \arrow[u, hook] \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta/2}(p_1,p_2,\cdots,p_n) \arrow[u, hook]
\end{tikzcd}
$$

$$\simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq$$

$$F_{\alpha_0} \subseteq \{p_1\}^{\delta/2} \subseteq \{p_1\}^{2\delta} \subseteq F_{\alpha_1} \subseteq \cdots\cdots \subseteq F_{\alpha_{n-1}} \subseteq \{p_1,\cdots,p_n\}^{\delta/2} \subseteq \{p_1,\cdots,p_n\}^{2\delta}$$

Pairs of Rips Complexes

$$\begin{tikzcd}
M_{\alpha-2c\delta} \arrow[rrr] \arrow[dr] & & & M_{\beta+2c\delta} \\
& N_\alpha \arrow[r] & N_\beta \arrow[ur] &
\end{tikzcd}$$

$$\begin{tikzcd}
\mathcal{N}^{2\delta}(p_1) \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1,p_2) \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1,p_2,\cdots,p_n) \\
\mathcal{R}^{2\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{R}^{2\delta}(p_1,p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{2\delta}(p_1,p_2,\cdots,p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{N}^{\delta}(p_1,p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta}(p_1,p_2,\cdots,p_n) \arrow[u, hook] \\
\mathcal{R}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{R}^{\delta}(p_1,p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{\delta}(p_1,p_2,\cdots,p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta/2}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{N}^{\delta/2}(p_1,p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta/2}(p_1,p_2,\cdots,p_n) \arrow[u, hook]
\end{tikzcd}$$

$$\simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq$$

$$F_{\alpha_0} \subseteq \{p_1\}^{\delta/2} \subseteq \{p_1\}^{2\delta} \subseteq F_{\alpha_1} \subseteq \cdots\cdots \subseteq F_{\alpha_{n-1}} \subseteq \{p_1,\cdots,p_n\}^{\delta/2} \subseteq \{p_1,\cdots,p_n\}^{2\delta}$$

Pairs of Rips Complexes

$$
\begin{tikzcd}
{\color{blue}M_\alpha} \arrow[r] \arrow[dr] & {\color{blue}M_\beta} \arrow[dr] & \\
& {\color{red}N_{\alpha+2c\delta}} \arrow[r] & {\color{red}N_{\beta+2c\delta}}
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathcal{N}^{2\delta}(p_1) \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1,p_2) \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1,p_2,\cdots,p_n) \\
{\color{red}\mathcal{R}^{2\delta}(p_1)} \arrow[r, hook, red] \arrow[u, hook] & {\color{red}\mathcal{R}^{2\delta}(p_1,p_2)} \arrow[r, hook, red] \arrow[u, hook] & {\color{red}\cdots} \arrow[r, hook, red] & {\color{red}\mathcal{R}^{2\delta}(p_1,p_2,\cdots,p_n)} \arrow[u, hook] \\
\mathcal{N}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook, red] & \mathcal{N}^{\delta}(p_1,p_2) \arrow[r, hook] \arrow[u, hook, red] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta}(p_1,p_2,\cdots,p_n) \arrow[u, hook, red] \\
{\color{red}\mathcal{R}^{\delta}(p_1)} \arrow[r, hook, red] \arrow[u, hook, red] & {\color{red}\mathcal{R}^{\delta}(p_1,p_2)} \arrow[r, hook, red] \arrow[u, hook, red] & {\color{red}\cdots} \arrow[r, hook, red] & {\color{red}\mathcal{R}^{\delta}(p_1,p_2,\cdots,p_n)} \arrow[u, hook, red] \\
\mathcal{N}^{\delta/2}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{N}^{\delta/2}(p_1,p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta/2}(p_1,p_2,\cdots,p_n) \arrow[u, hook]
\end{tikzcd}
$$

$$\simeq \qquad \simeq \qquad \simeq \qquad \simeq \qquad \simeq \qquad \simeq$$

$${\color{blue}F_{\alpha_0}} \subseteq \{p_1\}^{\delta/2} \subseteq \{p_1\}^{2\delta} \subseteq {\color{blue}F_{\alpha_1}} \subseteq \cdots\cdots \subseteq {\color{blue}F_{\alpha_{n-1}}} \subseteq \{p_1,\cdots,p_n\}^{\delta/2} \subseteq \{p_1,\cdots,p_n\}^{2\delta}$$

Pairs of Rips Complexes

$$\begin{tikzcd}
 & M_{\alpha+2c\delta} \arrow[r] & M_{\beta+2c\delta} \\
N_\alpha \arrow[ur] \arrow[r] & N_\beta \arrow[ur] &
\end{tikzcd}$$

$$\begin{tikzcd}
\mathcal{N}^{2\delta}(p_1) \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1, p_2) \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{2\delta}(p_1, p_2, \cdots, p_n) \\
\mathcal{R}^{2\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{R}^{2\delta}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{2\delta}(p_1, p_2, \cdots, p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{N}^{\delta}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta}(p_1, p_2, \cdots, p_n) \arrow[u, hook] \\
\mathcal{R}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{R}^{\delta}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{\delta}(p_1, p_2, \cdots, p_n) \arrow[u, hook] \\
\mathcal{N}^{\delta/2}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{N}^{\delta/2}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{N}^{\delta/2}(p_1, p_2, \cdots, p_n) \arrow[u, hook]
\end{tikzcd}$$

$$\simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq \quad \simeq$$

$$F_{\alpha_0} \subseteq \{p_1\}^{\delta/2} \subseteq \{p_1\}^{2\delta} \subseteq F_{\alpha_1} \subseteq \cdots\cdots \subseteq F_{\alpha_{n-1}} \subseteq \{p_1, \cdots, p_n\}^{\delta/2} \subseteq \{p_1, \cdots, p_n\}^{2\delta}$$

Pairs of Rips Complexes

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

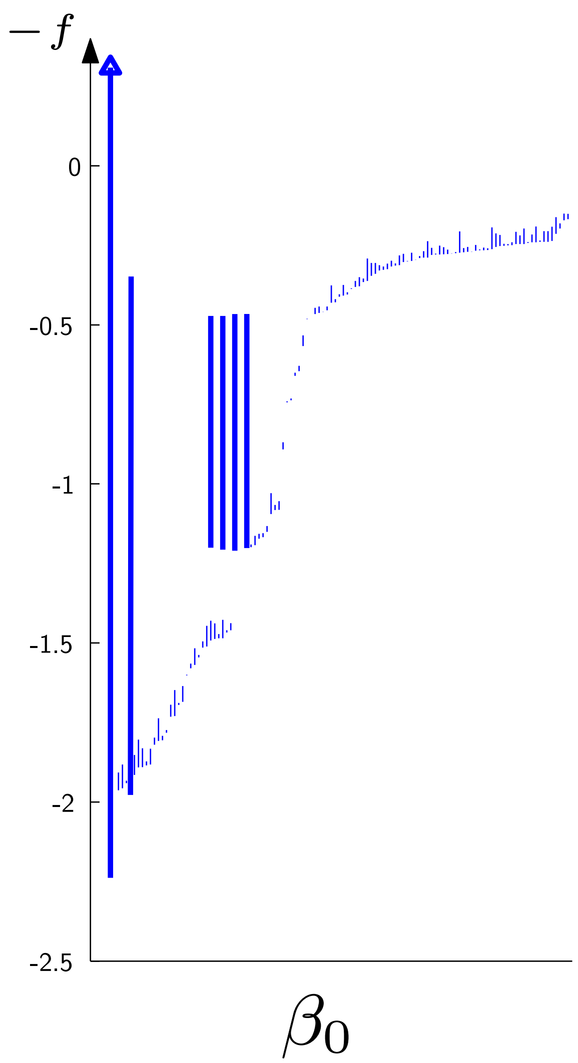
Guarantees: $\forall \delta \in [2\varepsilon, \frac{1}{2}\varrho_c(\mathbb{X}))$,

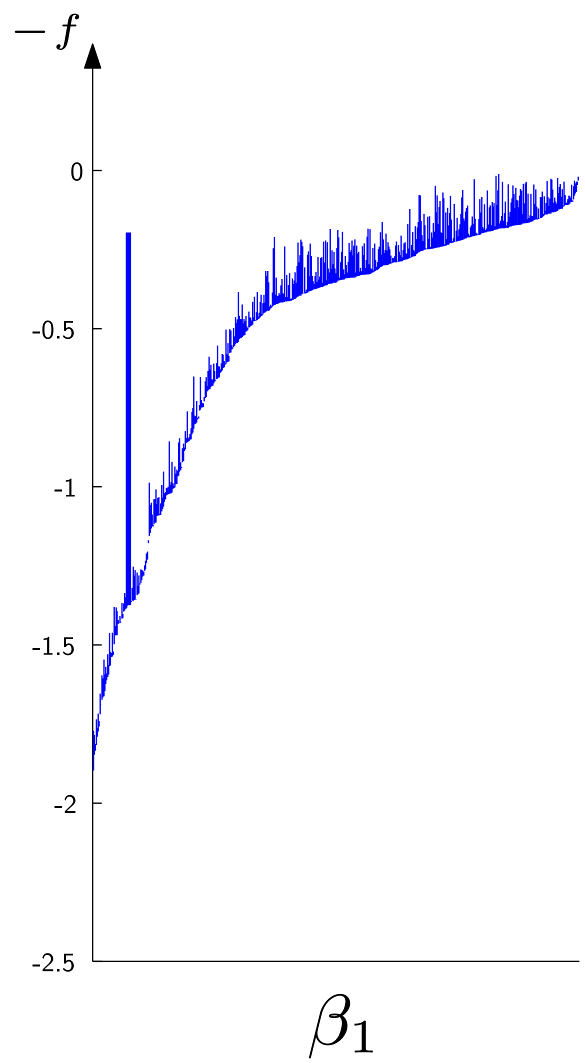
$\{H_k(F_\alpha)\}_{\alpha\in\mathbb{R}}$ and $\{\mathrm{im}\; H_k(\mathcal{R}^\delta(P_\alpha)) \to H_k(\mathcal{R}^{2\delta}(P_\alpha))\}_{\alpha\in\mathbb{R}}$ are $2c\delta$-*interleaved*:

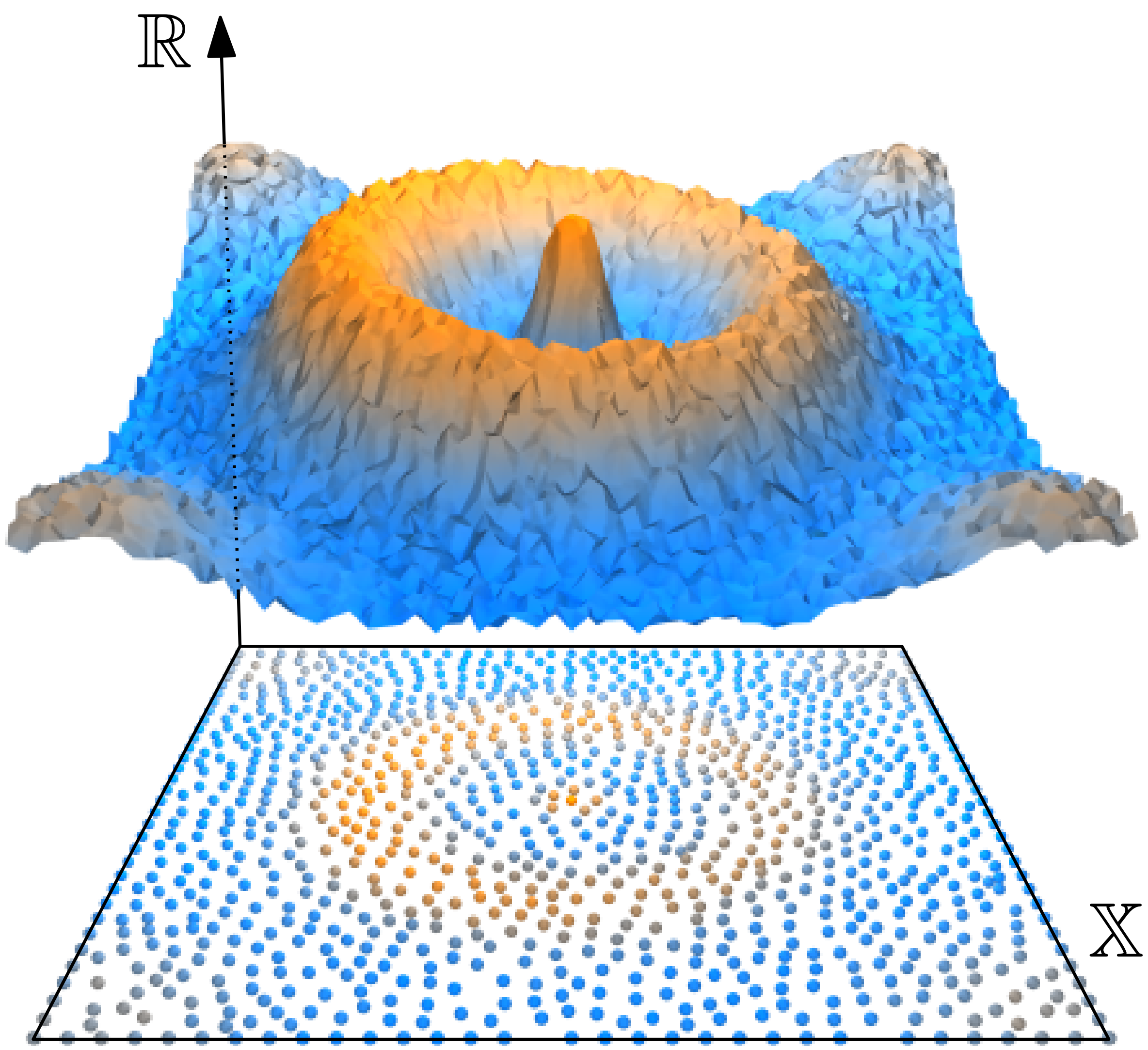
$\Downarrow$ [Chazal, Cohen-Steiner, Glisse, Guibas, O. '09]

their persistence diagrams are $2c\delta$-close.

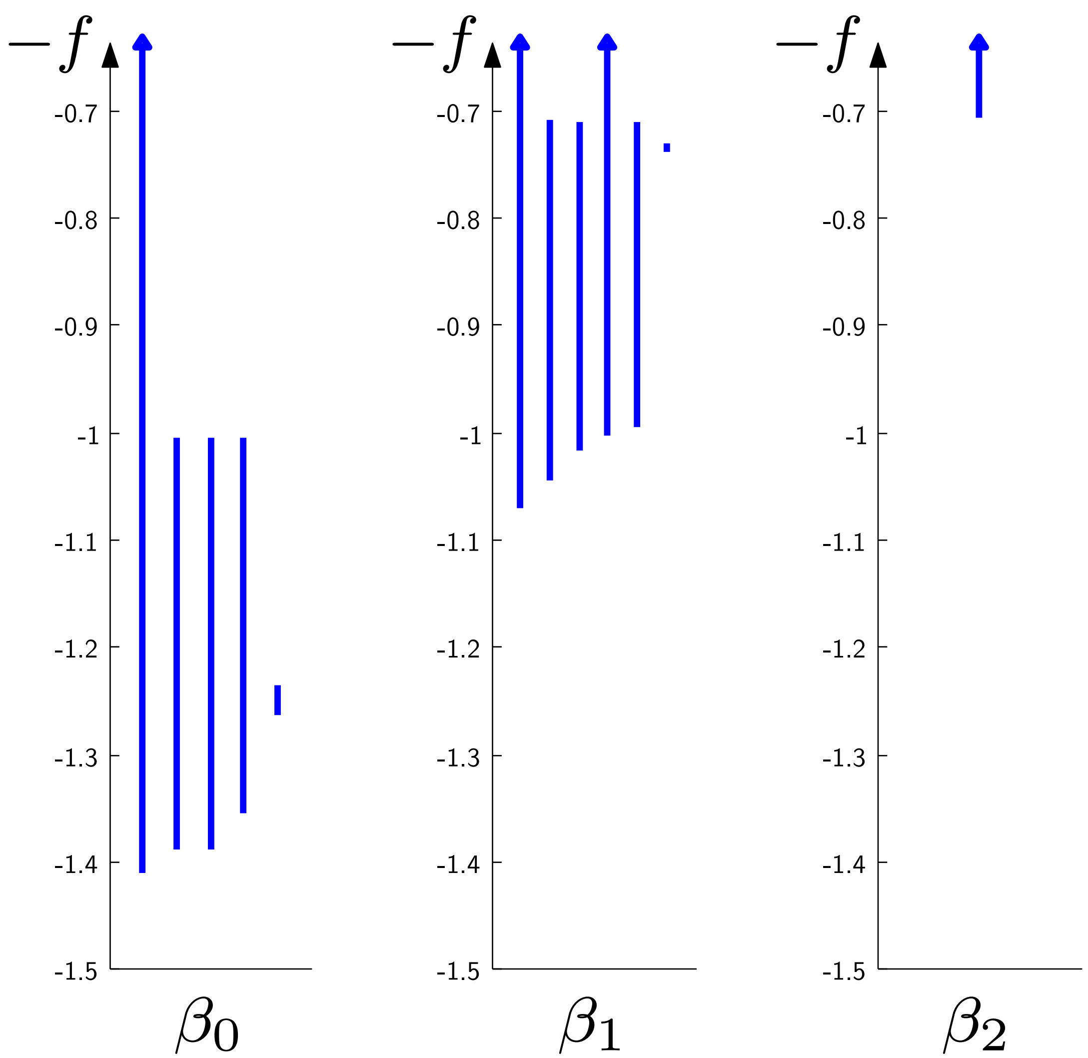
Some Results



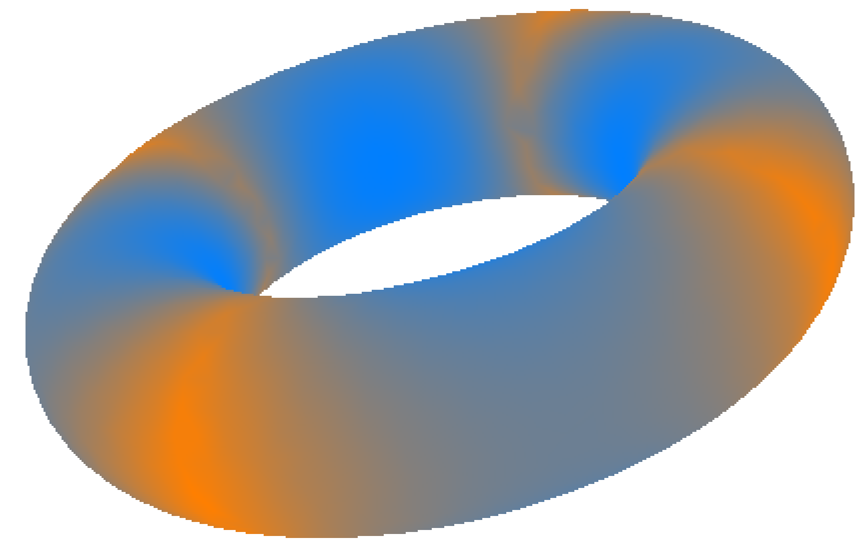




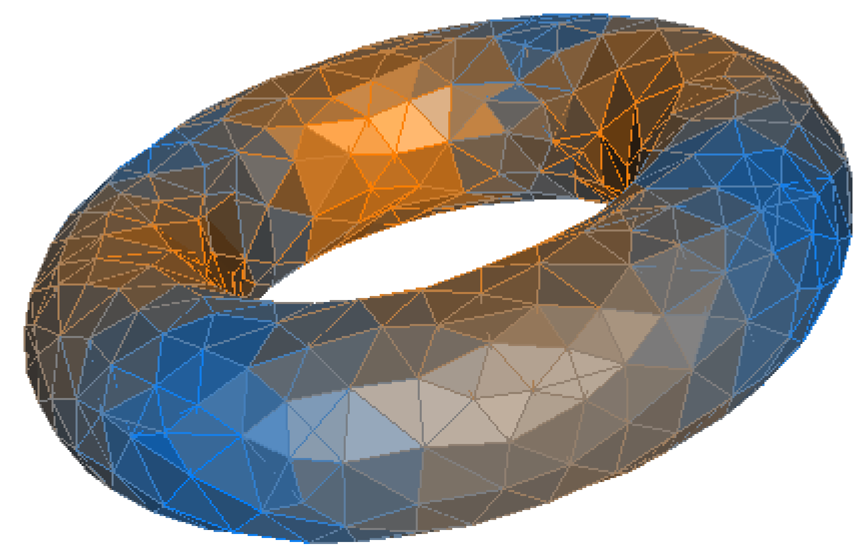
Some Results



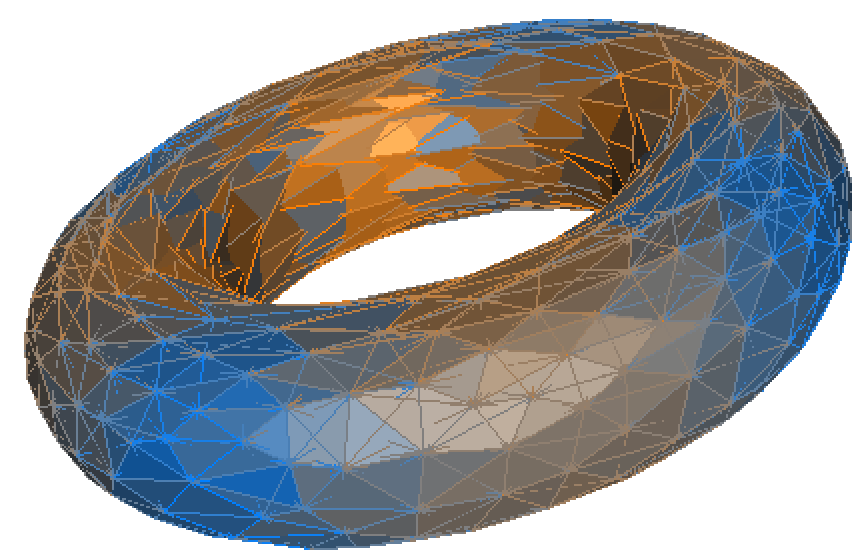
$(\mathbb{X}, f)$



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



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



Back to the Algorithm

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Algorithm: choose parameter $\delta \geq 0$

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

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

Back to the Algorithm

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Algorithm: choose parameter $\delta \geq 0$

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

$$\begin{tikzcd}
\mathcal{R}^{2\delta}(p_1) \arrow[r, hook] & \mathcal{R}^{2\delta}(p_1, p_2) \arrow[r, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{2\delta}(p_1, p_2, \cdots, p_n) \\
\mathcal{R}^{\delta}(p_1) \arrow[r, hook] \arrow[u, hook] & \mathcal{R}^{\delta}(p_1, p_2) \arrow[r, hook] \arrow[u, hook] & \cdots \arrow[r, hook] & \mathcal{R}^{\delta}(p_1, p_2, \cdots, p_n) \arrow[u, hook]
\end{tikzcd}$$

Obs.: At 0th homology level, the vertical maps induce surjective homomorphisms

Back to the Algorithm

Assume $\mathbb{X}$ is a Riemannian manifold, $f : \mathbb{X} \to \mathbb{R}$ is c-Lipschitz,
P is an ε-sample of $\mathbb{X}$ for some (unknown) $0 < \varepsilon < \frac{1}{4}\varrho_c(\mathbb{X})$.

Algorithm: choose parameter $\delta \geq 0$

1. Sort the data points such that $f(p_1) \leq f(p_2) \leq \cdots \leq f(p_n)$,
2. For $i = 1, \cdots, n$, build the 1-skeleton of $\mathcal{R}^{2\delta}(p_1, \cdots, p_i)$,
3. Apply the persistence algorithm to the filtration

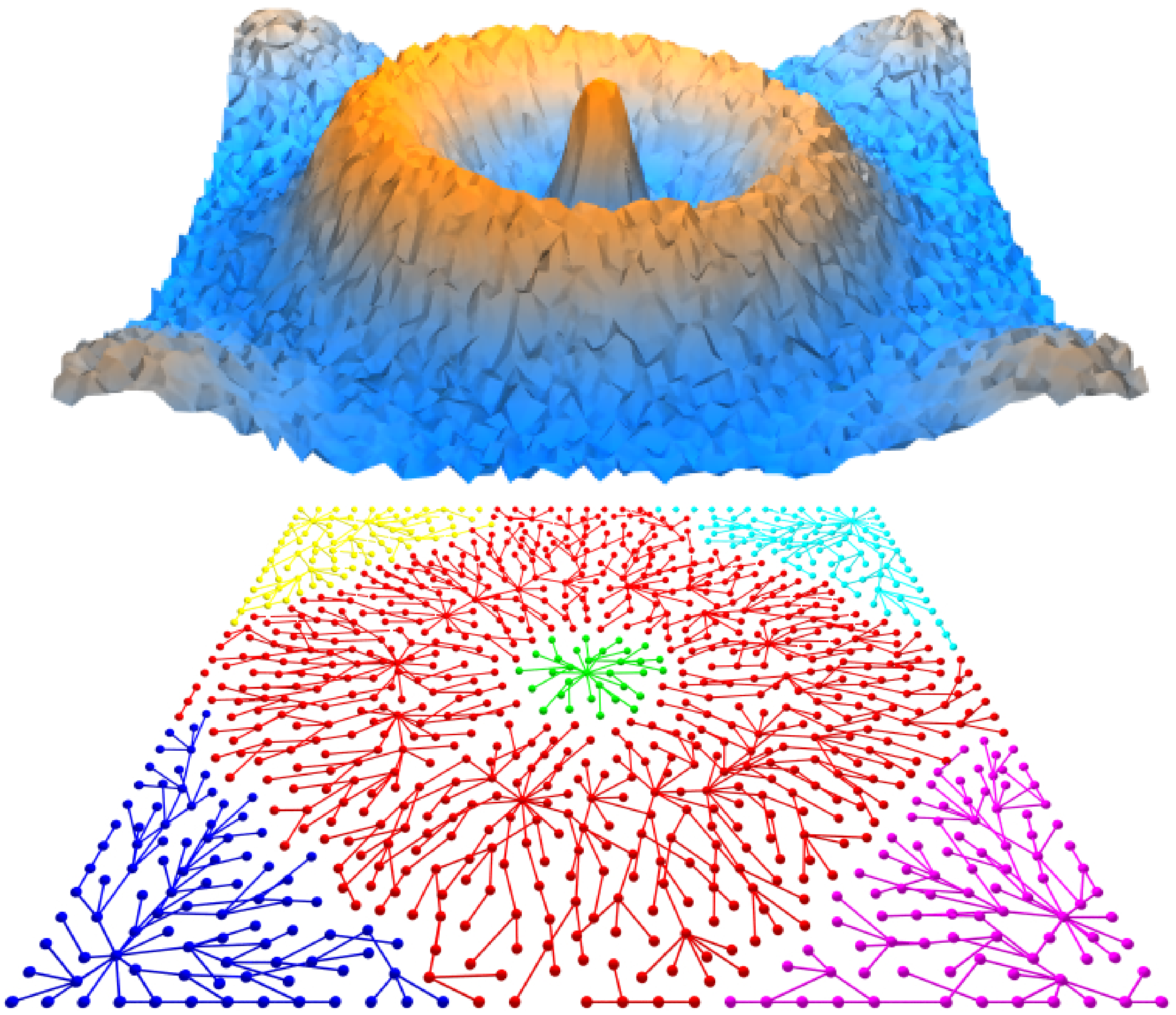
$$\mathcal{R}^{2\delta}(p_1) \hookrightarrow \mathcal{R}^{2\delta}(p_1, p_2) \hookrightarrow \cdots \hookrightarrow \mathcal{R}^{2\delta}(p_1, p_2, \cdots, p_n)$$

Obs.: At 0th homology level, the vertical maps induce surjective homomorphisms

Basins of Attraction

Goal: approximate basins of attraction of significant peaks of f

$\rightarrow$ cluster the input point cloud P

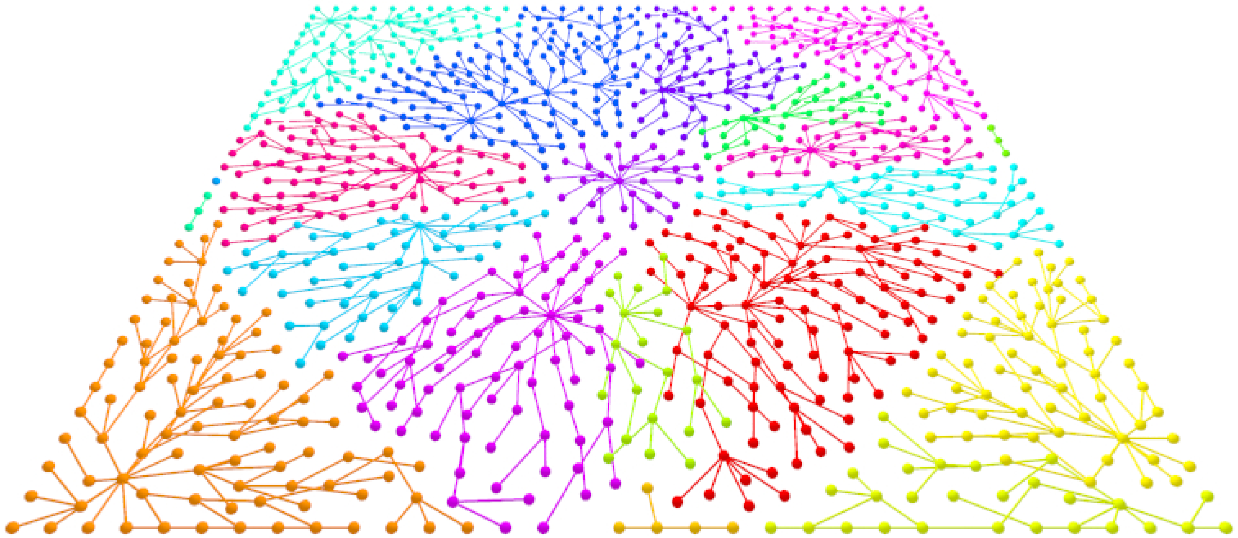


Basins of Attraction

Goal: approximate basins of attraction of significant peaks of f

$\rightarrow$ cluster the input point cloud P

Approach:



- simulate effect of ∇f by connecting vertex to highest neighbor in Rips graph
$\rightarrow$ gives a forest, the roots of which are local maxima of f in the graph

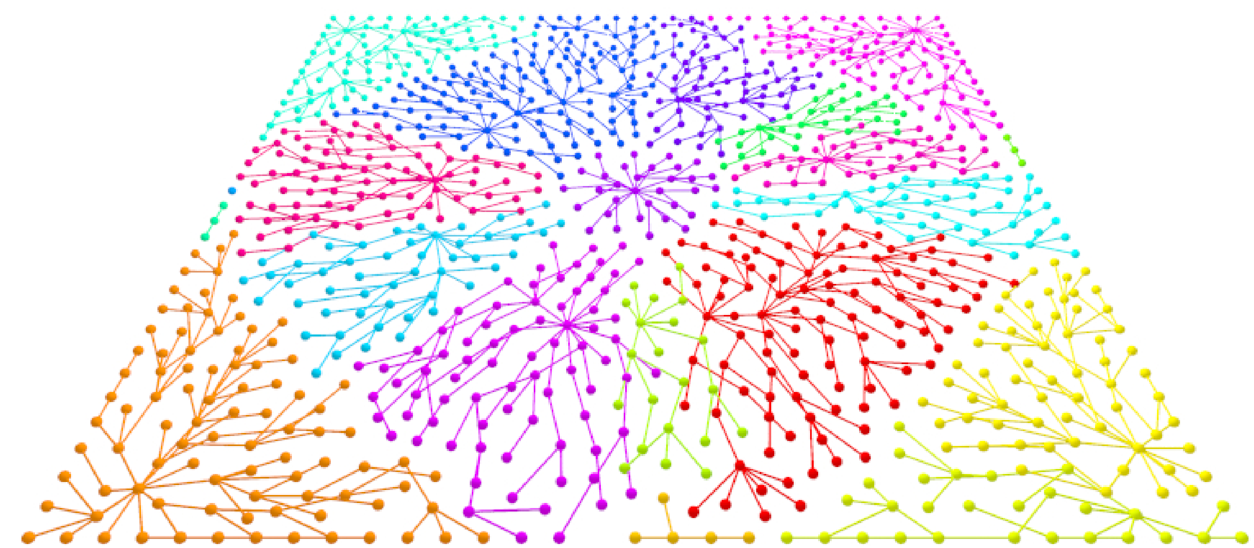
Basins of Attraction

Goal: approximate basins of attraction of significant peaks of f

$\longrightarrow$ cluster the input point cloud P

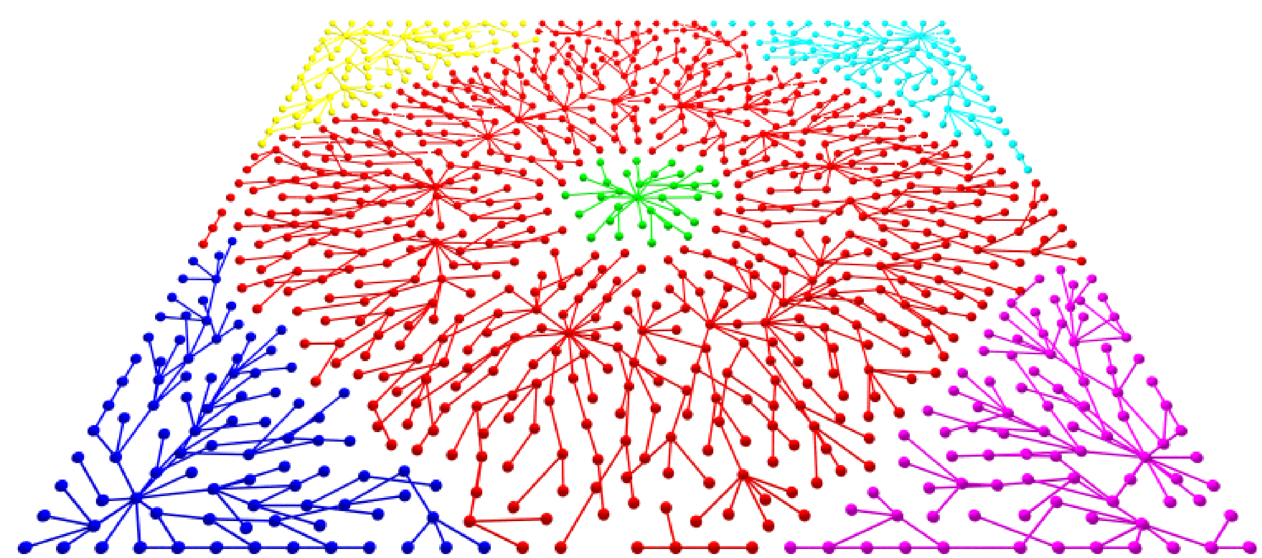
Approach:

• simulate effect of ∇f by connecting vertex to highest neighbor in Rips graph
$\longrightarrow$ gives a forest, the roots of which are local maxima of f in the graph



• apply 0-dimensional algorithm to $-f$
$\longrightarrow$ clusters are merged by persistence algorithm (union-find data structure)

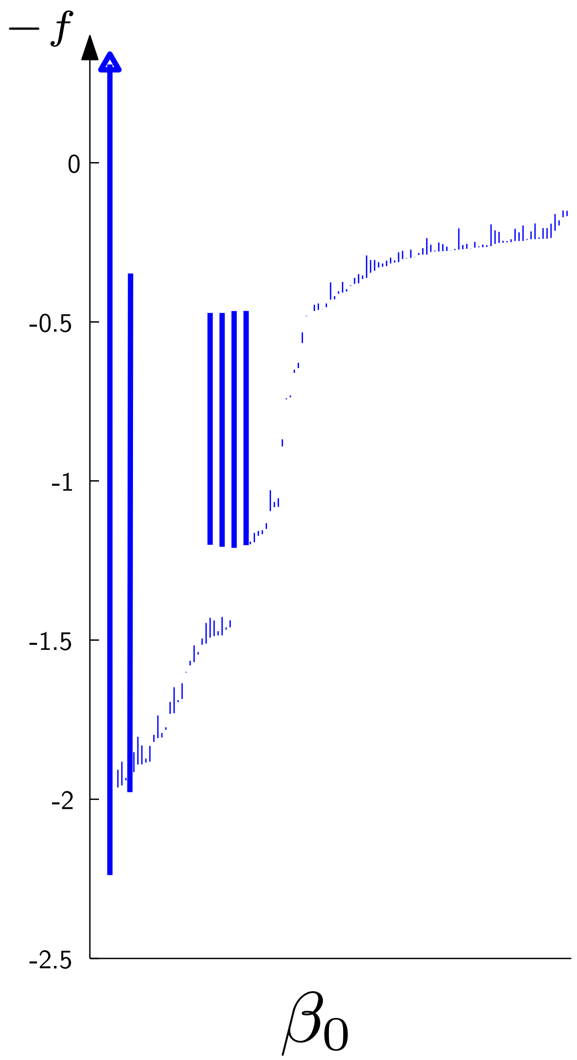
• do not merge clusters that are more persistent than a given threshold τ

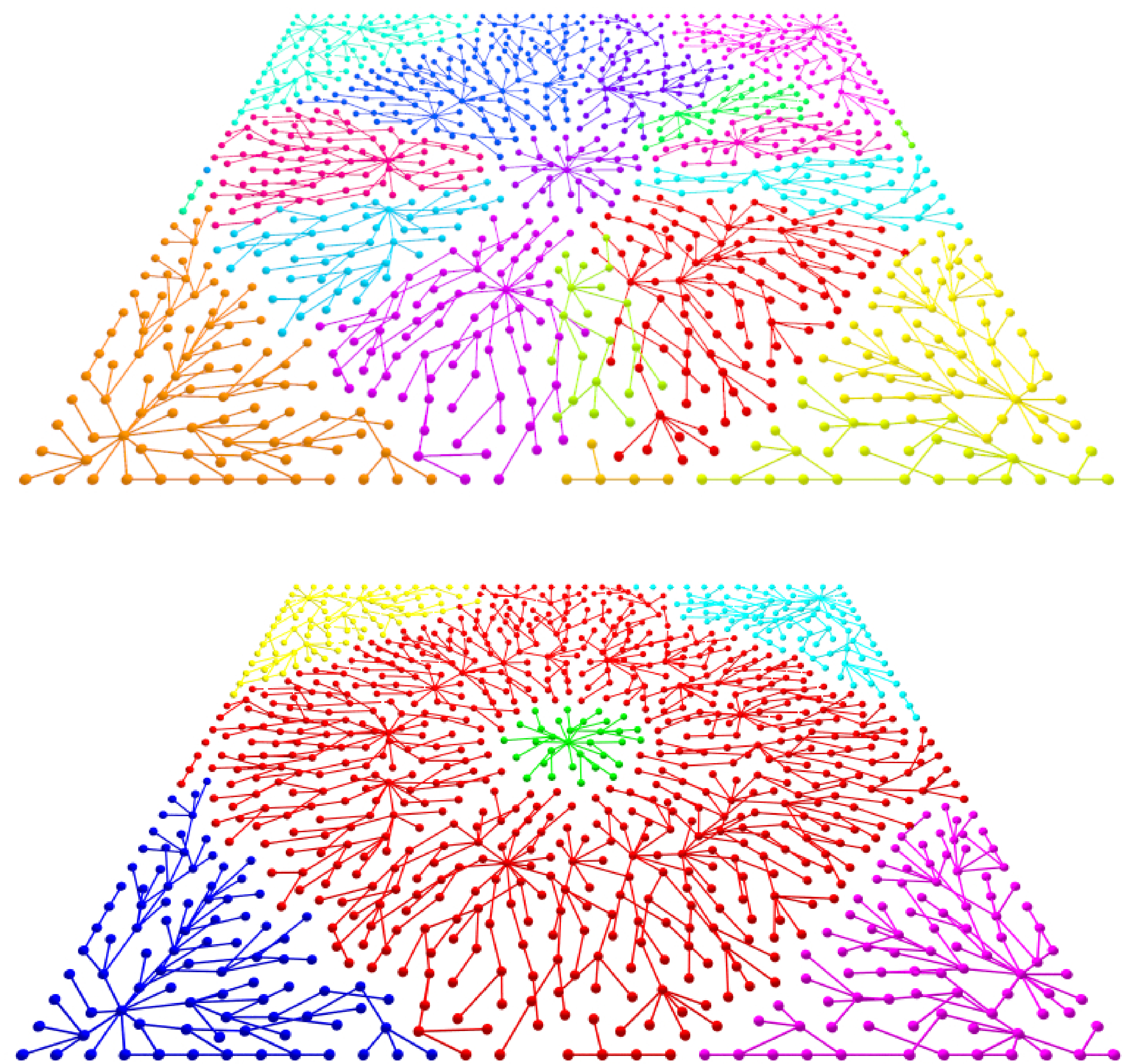


Basins of Attraction

Goal: approximate basins of attraction of significant peaks of f

$\longrightarrow$ cluster the input point cloud P





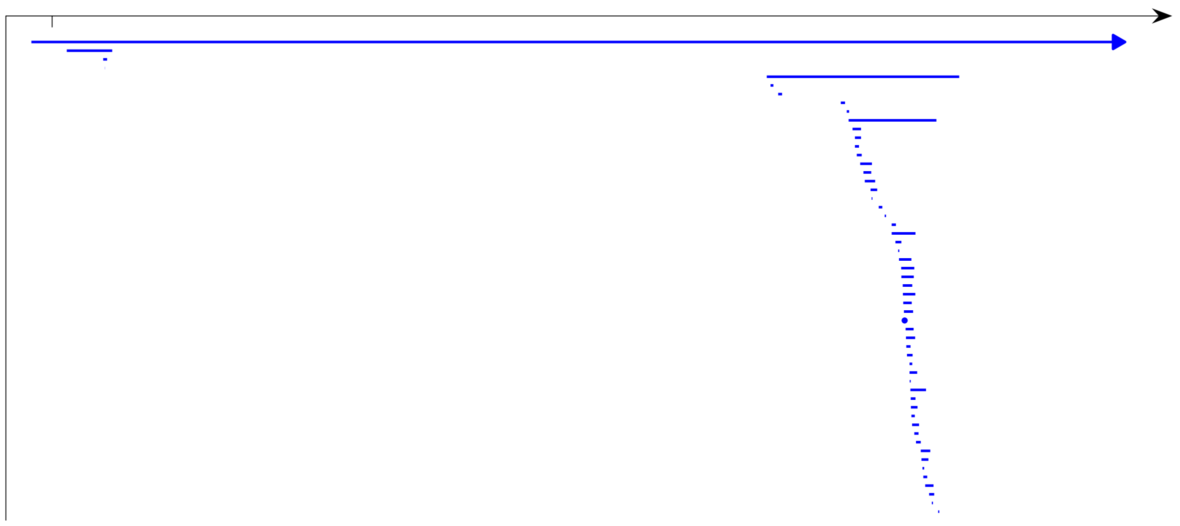
$\longrightarrow$ the user can use the 0-dim. persistence diagram to choose τ (two runs)

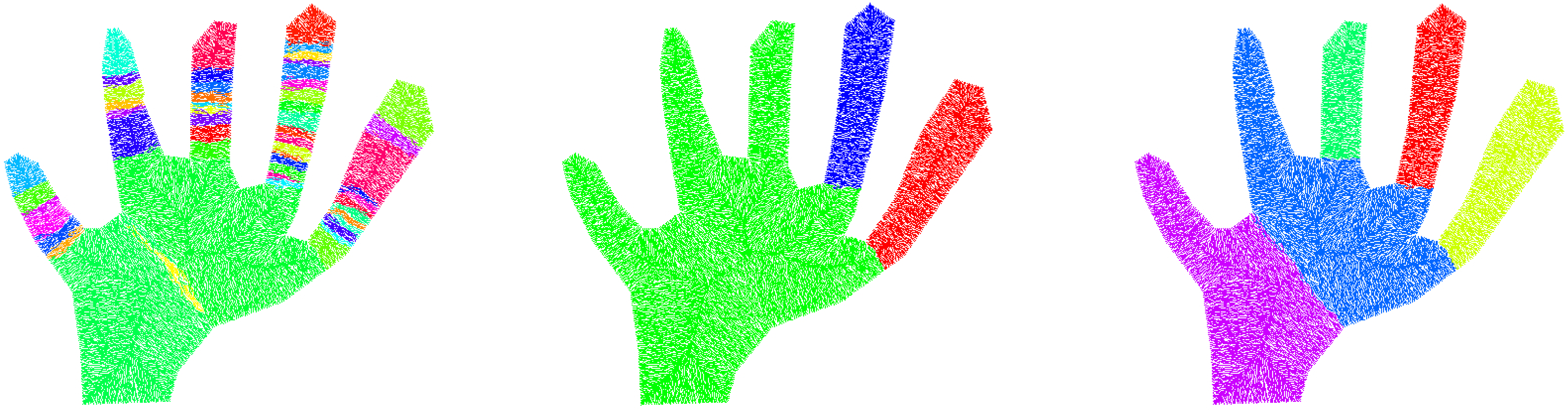
Application to Shape Segmentation

- Goal: partition a sampled shape into its most *natural parts* (ill-posed)

- Problem is cast into the one of finding a *good* segmentation function

$\longrightarrow$ distance ot the boundary:



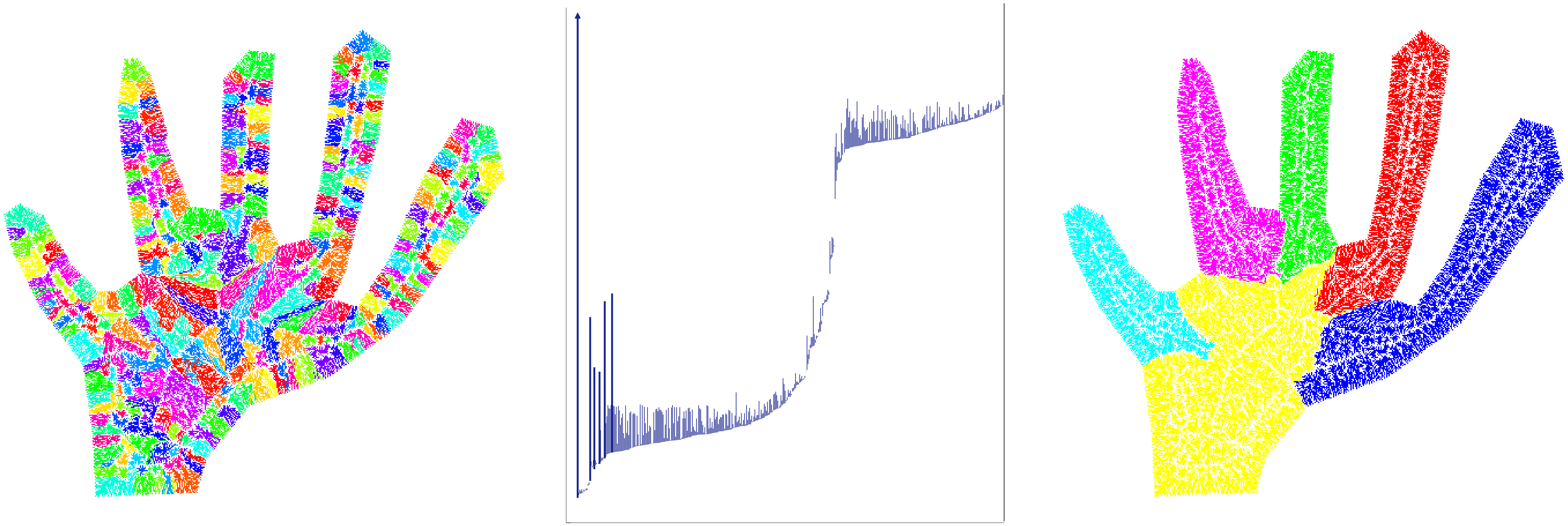


Application to Shape Segmentation

- Goal: partition a sampled shape into its most *natural parts* (ill-posed)

- Problem is cast into the one of finding a *good* segmentation function

$\longrightarrow$ normalized diameter of nearest boundary points:

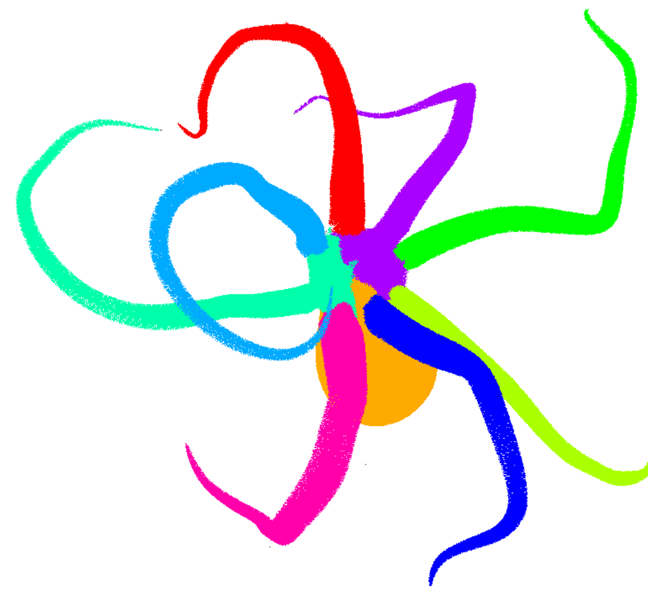


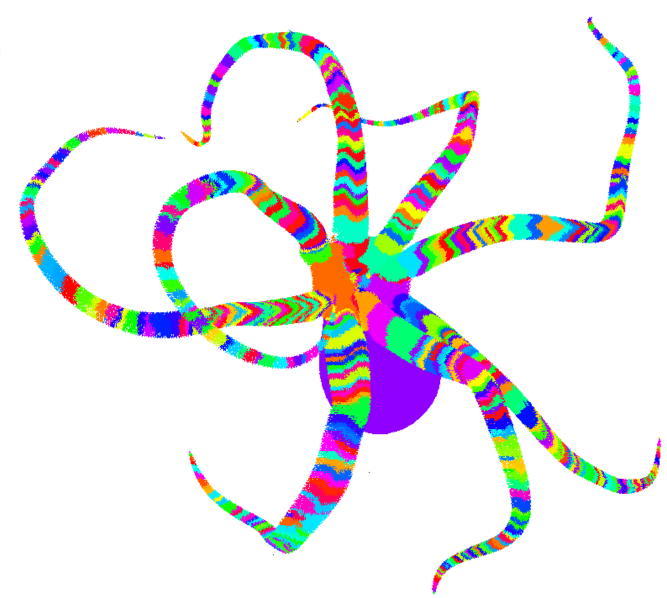
Application to Shape Segmentation

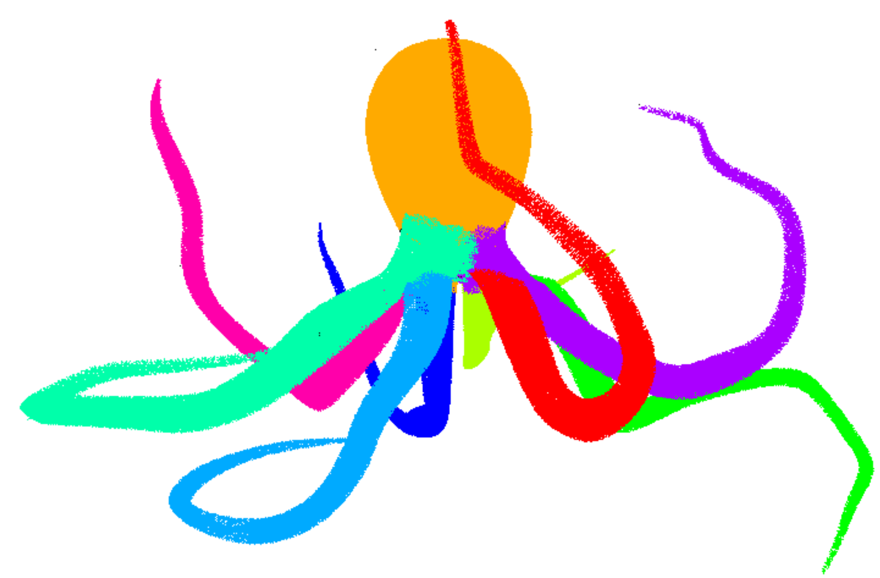
- Goal: partition a sampled shape into its most *natural parts* (ill-posed)

- Problem is cast into the one of finding a *good* segmentation function

$\longrightarrow$ an example in 3D:



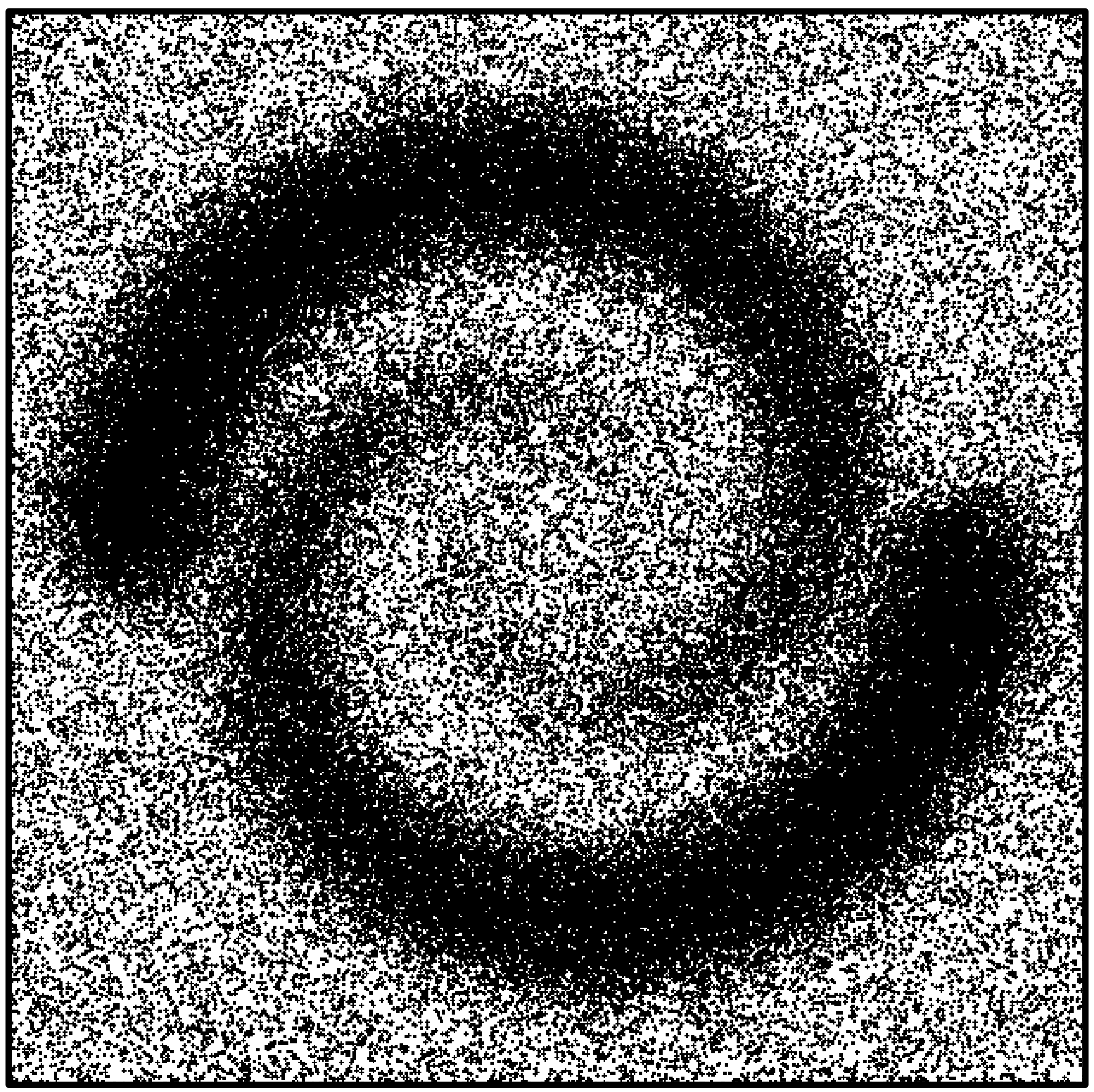


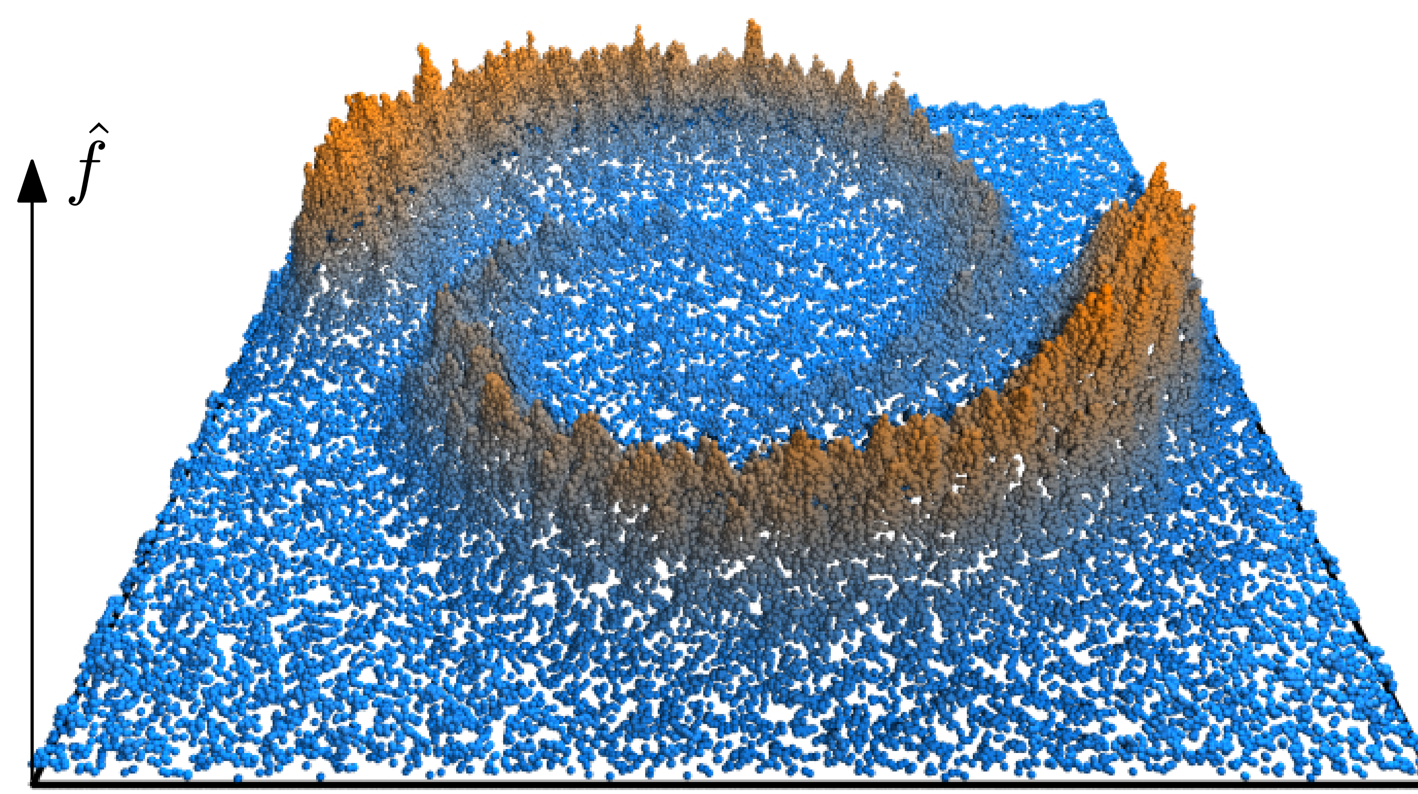


Application to Clustering

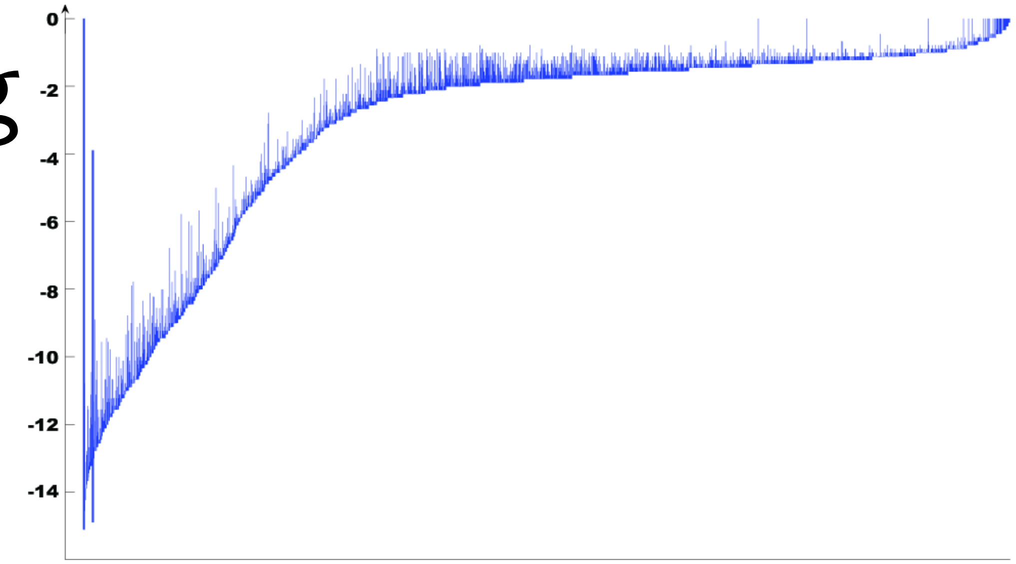
Input: $\mathbb{X} = [0,1]^2$; $|P| = 100,000$;

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



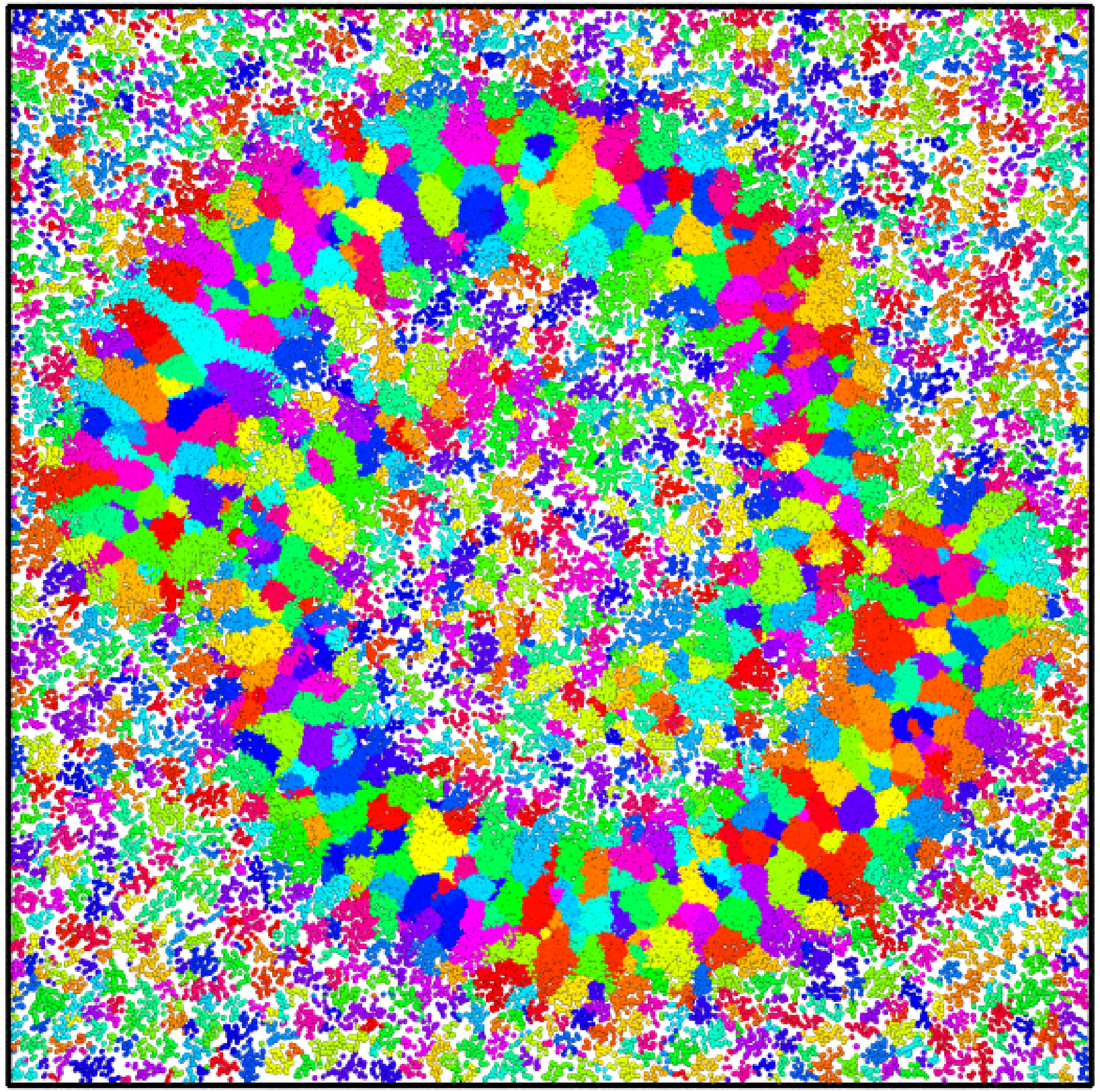


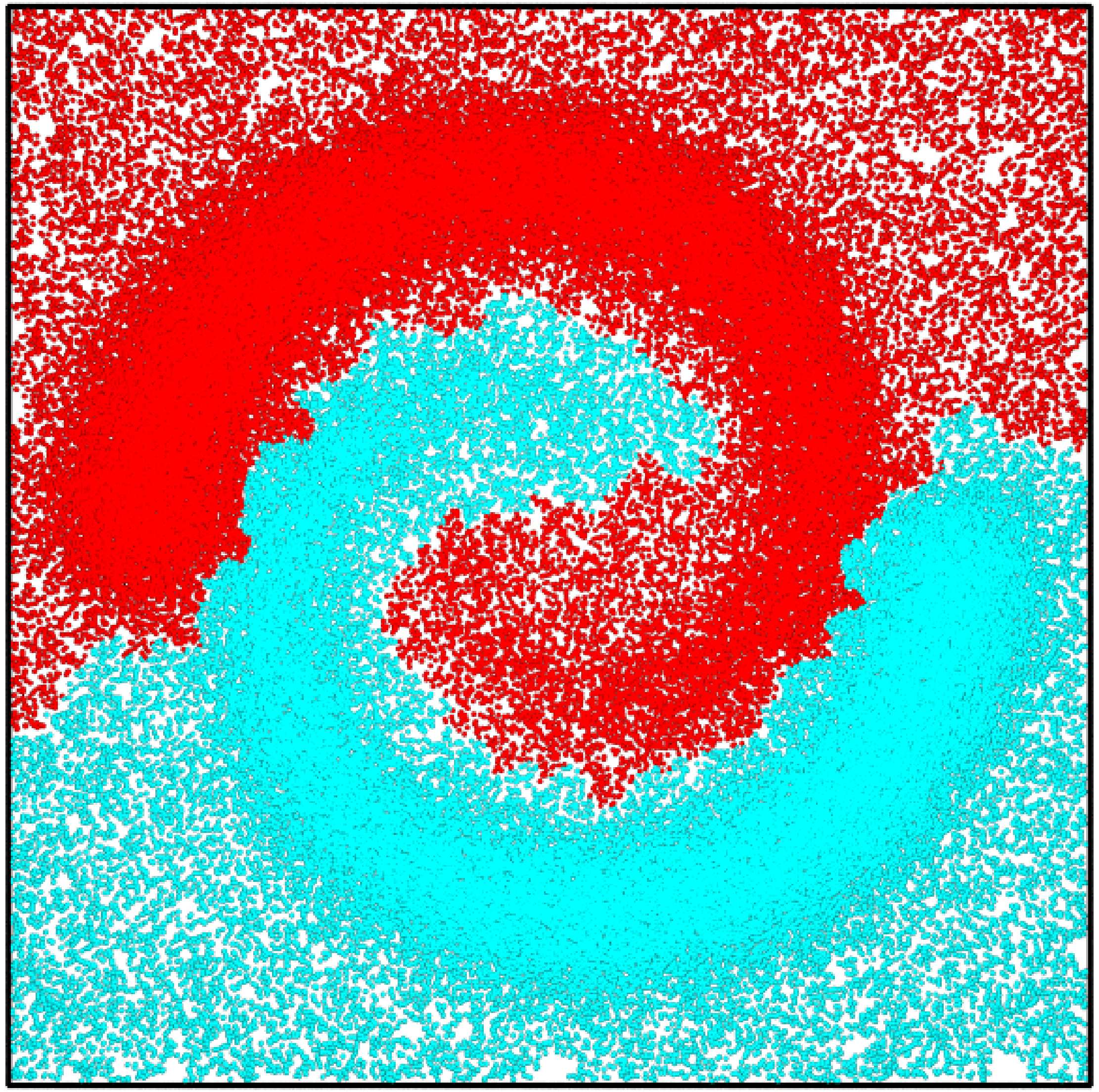
Application to Clustering



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

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

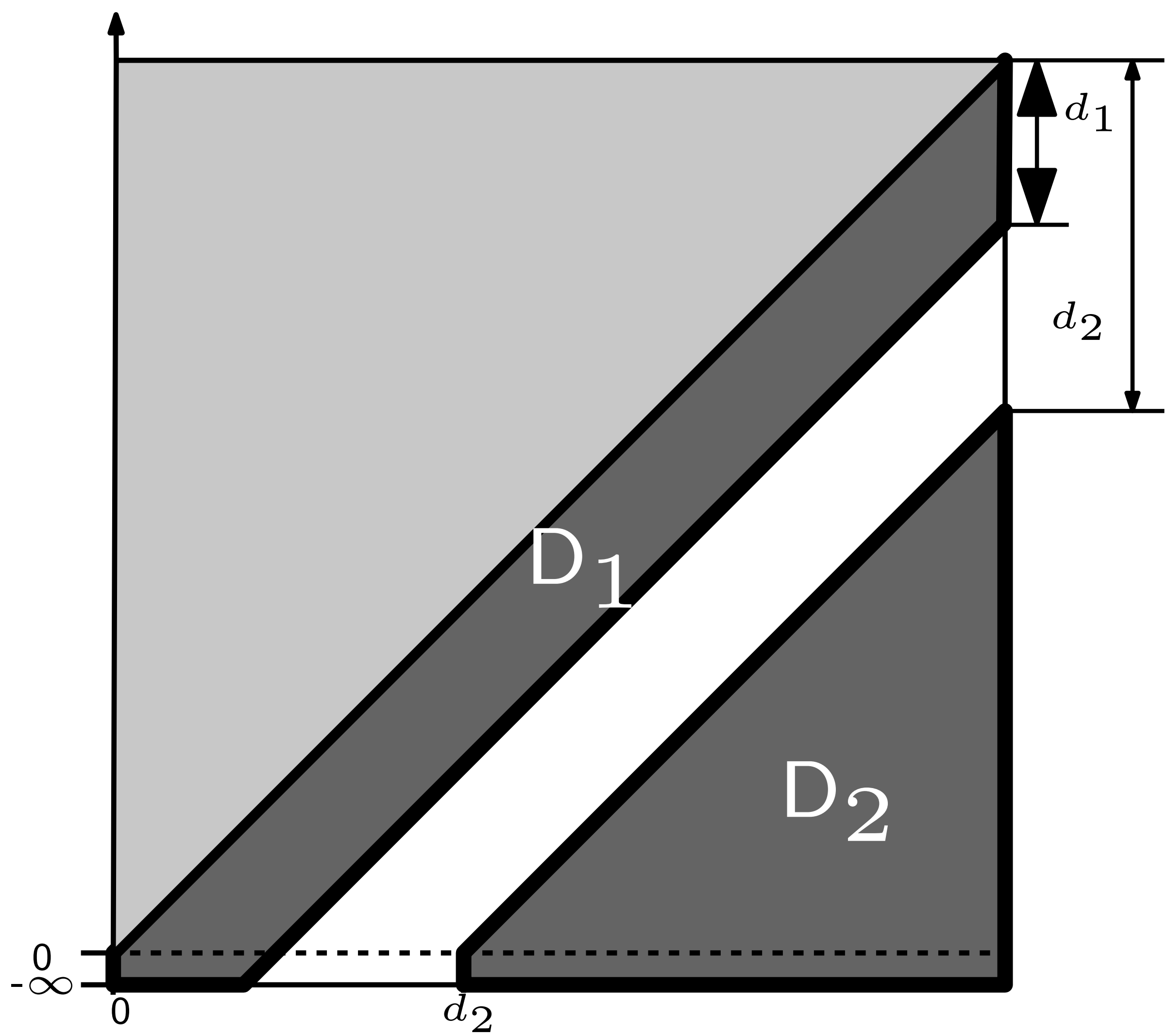




Application to Clustering

Main challenges:

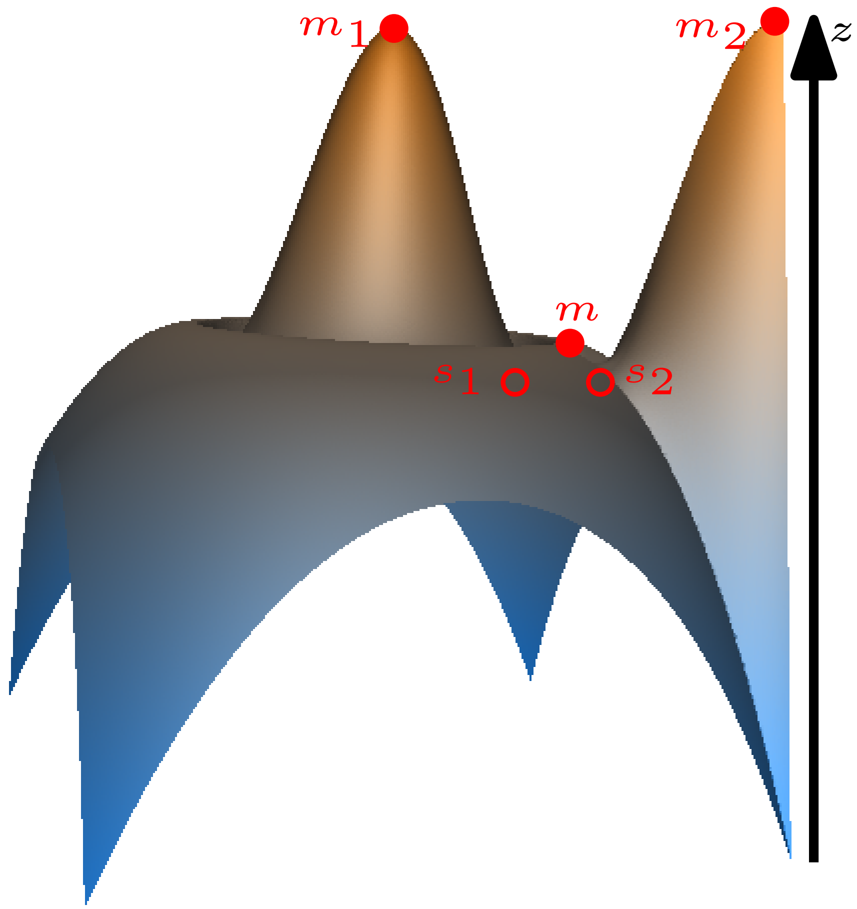
- finding the *correct* number of clusters (separation theorem)

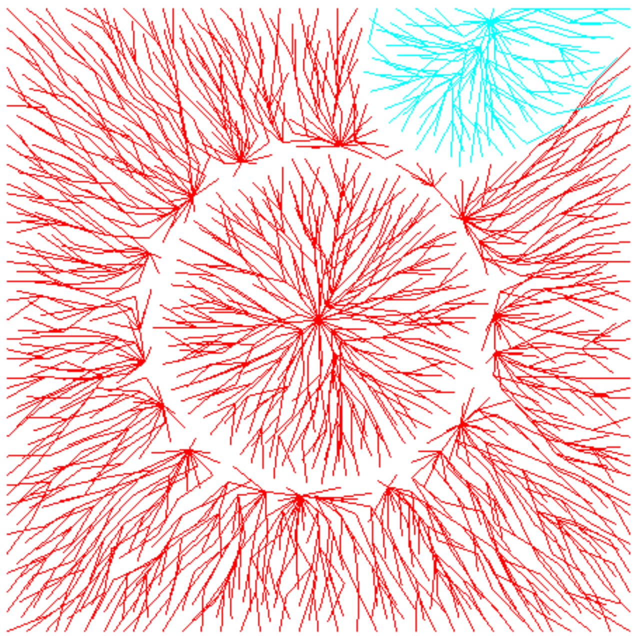


Application to Clustering

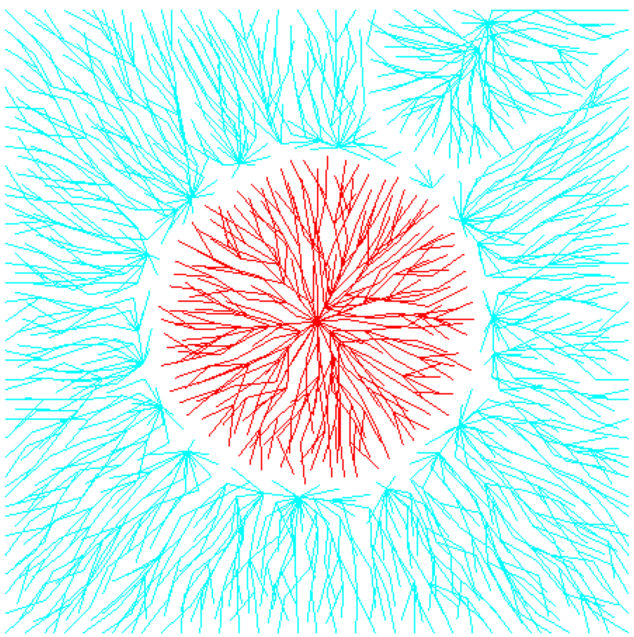
Main challenges:

- finding the *correct* number of clusters (separation theorem)
- approximating the basins of attraction (partial approximation result)





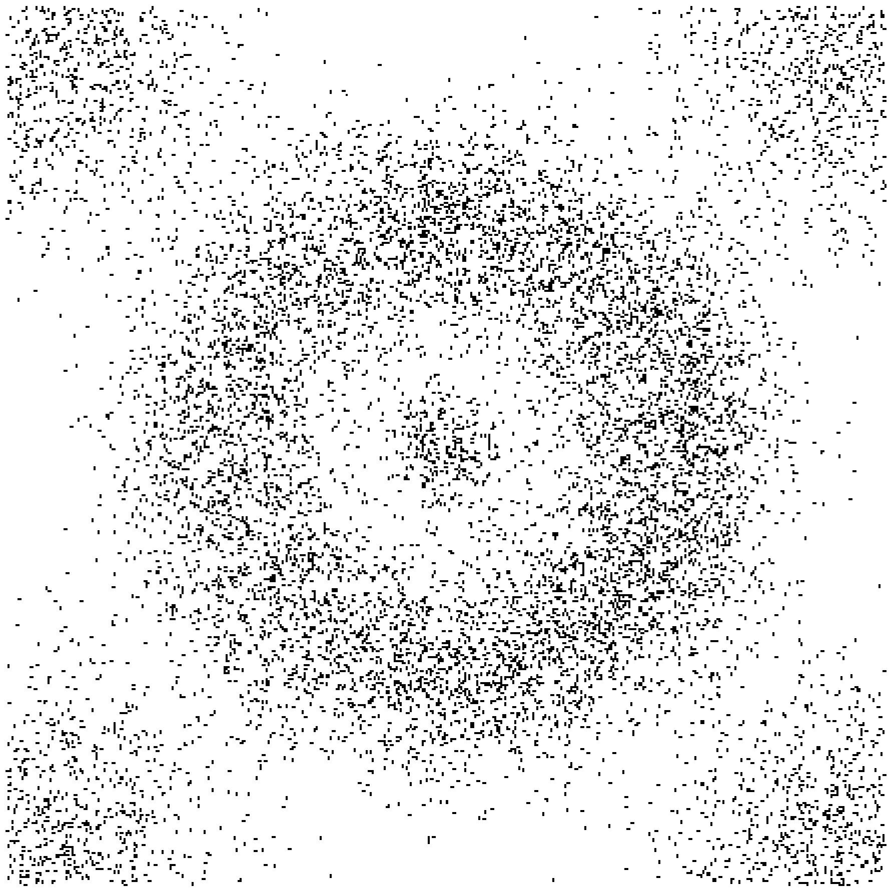
vs.



Application to Clustering

Main challenges:

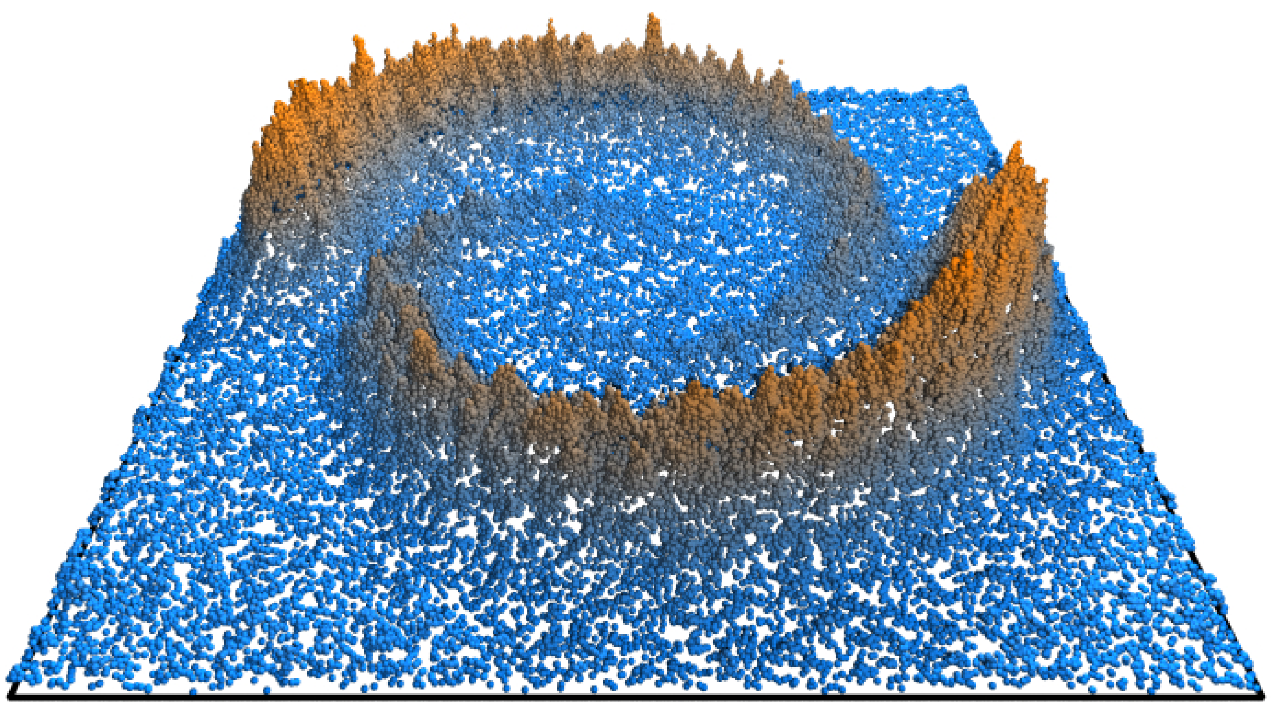
- finding the *correct* number of clusters (separation theorem)
- approximating the basins of attraction (partial approximation result)
- $\mathbb{X}$ may be partially sampled (yet some superlevel-set is well-sampled)



Application to Clustering

Main challenges:

- finding the *correct* number of clusters (separation theorem)
- approximating the basins of attraction (partial approximation result)
- $\mathbb{X}$ may be partially sampled (yet some superlevel-set is well-sampled)
- density is only known through some noisy estimator (stability results)



Application to Clustering

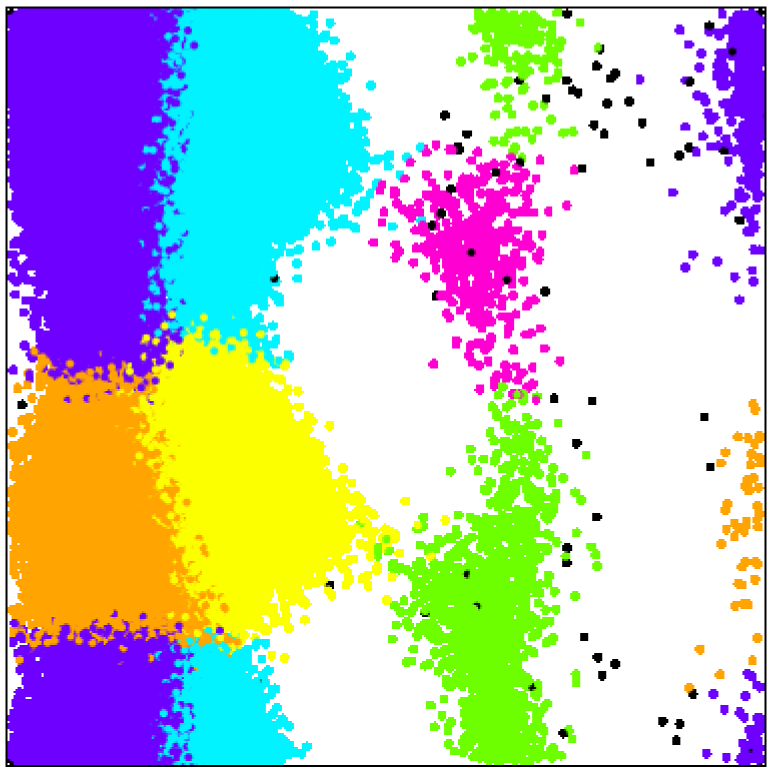
Main challenges:

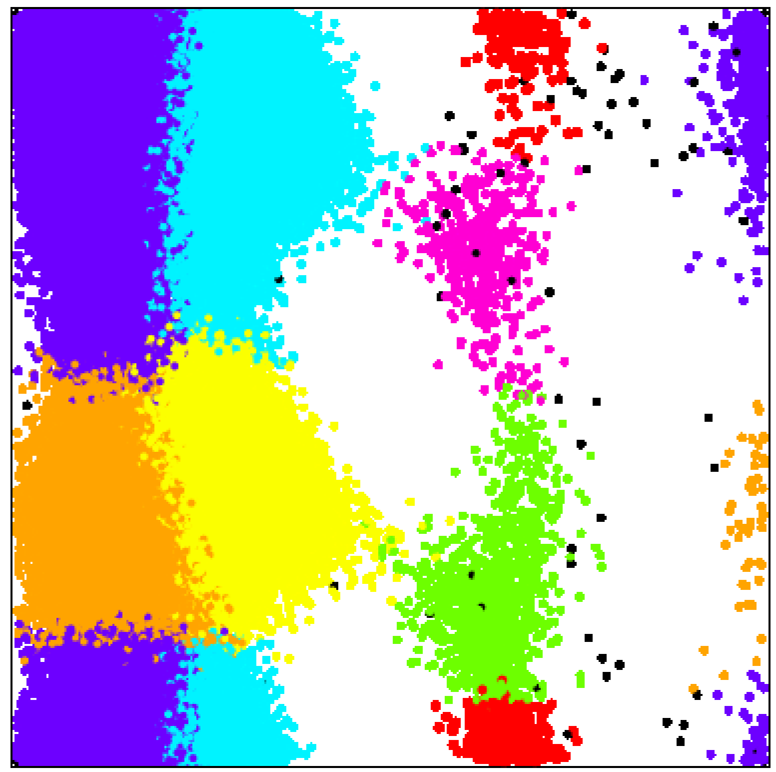
- finding the *correct* number of clusters (separation theorem)
- approximating the basins of attraction (partial approximation result)
- $\mathbb{X}$ may be partially sampled (yet some superlevel-set is well-sampled)
- density is only known through some noisy estimator (stability results)
- pairwise distances are often approximated (GH-stability)

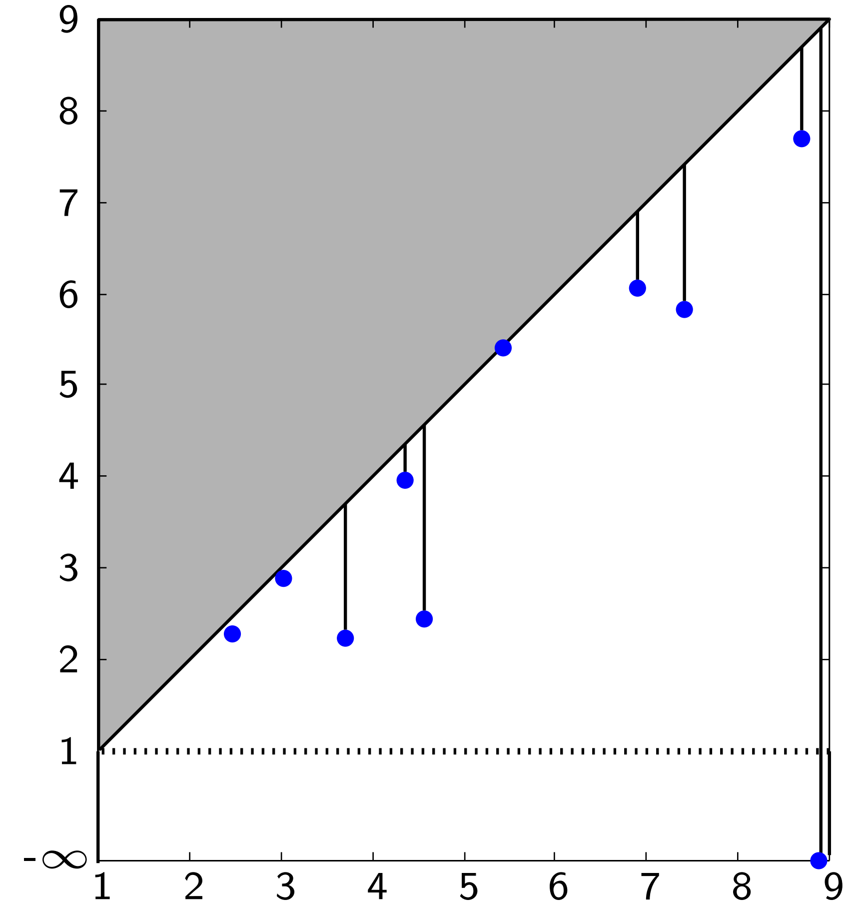
Application to Clustering

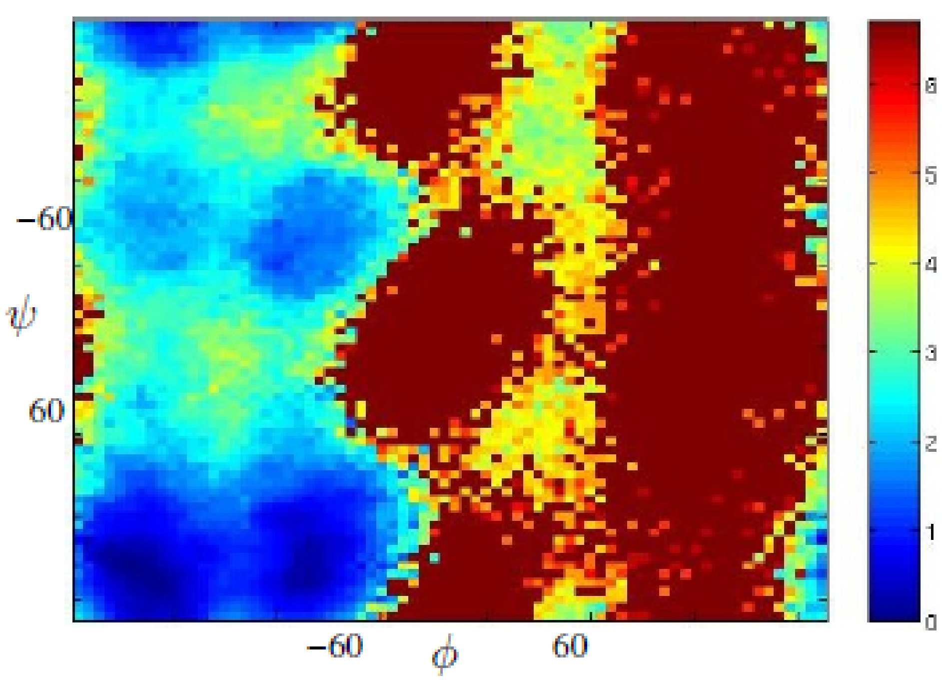
Back to the Alanine-dipeptide dataset:

- clustering performed in R^{21}; persistence diagram plotted on a log/log scale





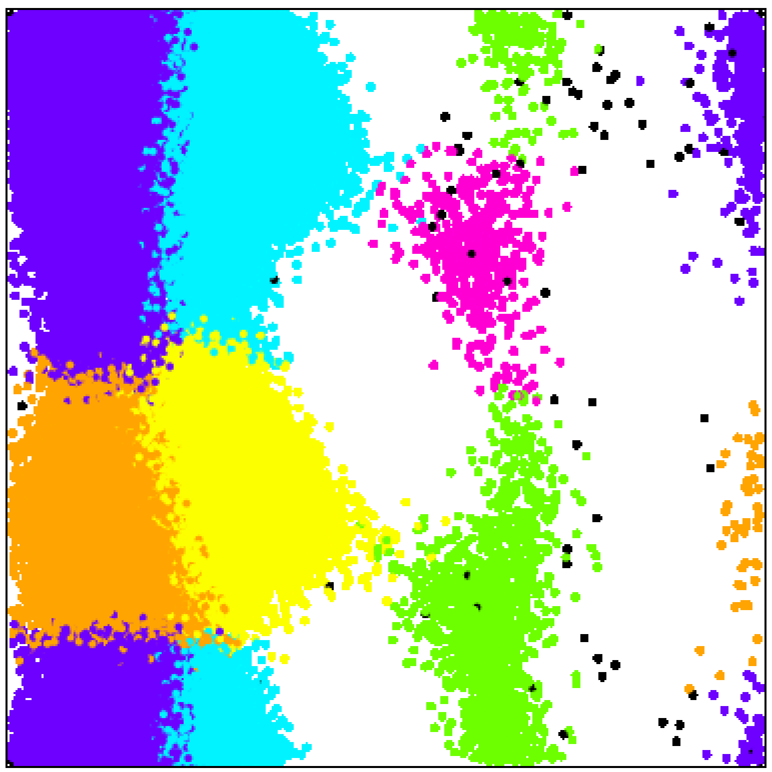


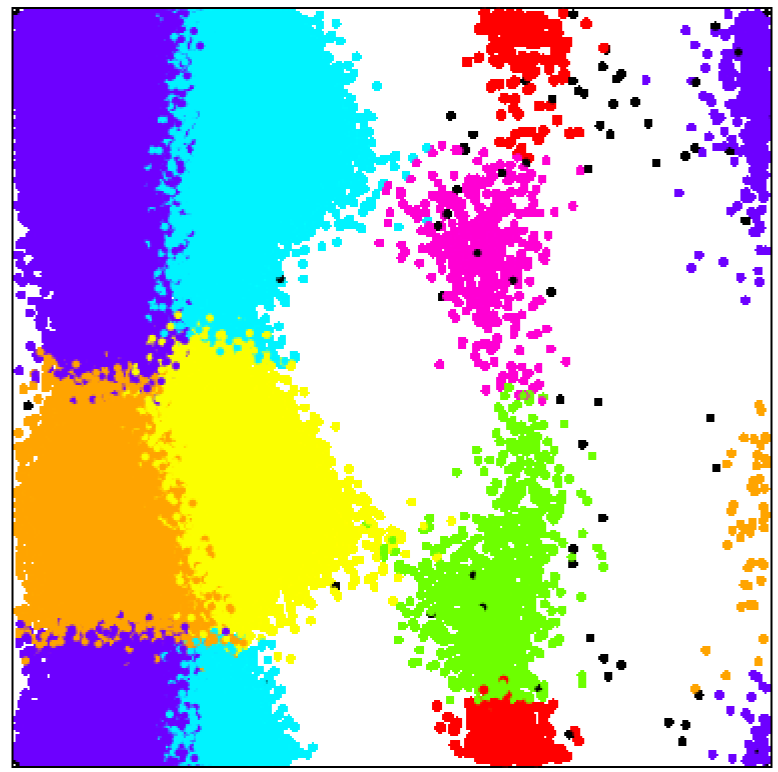


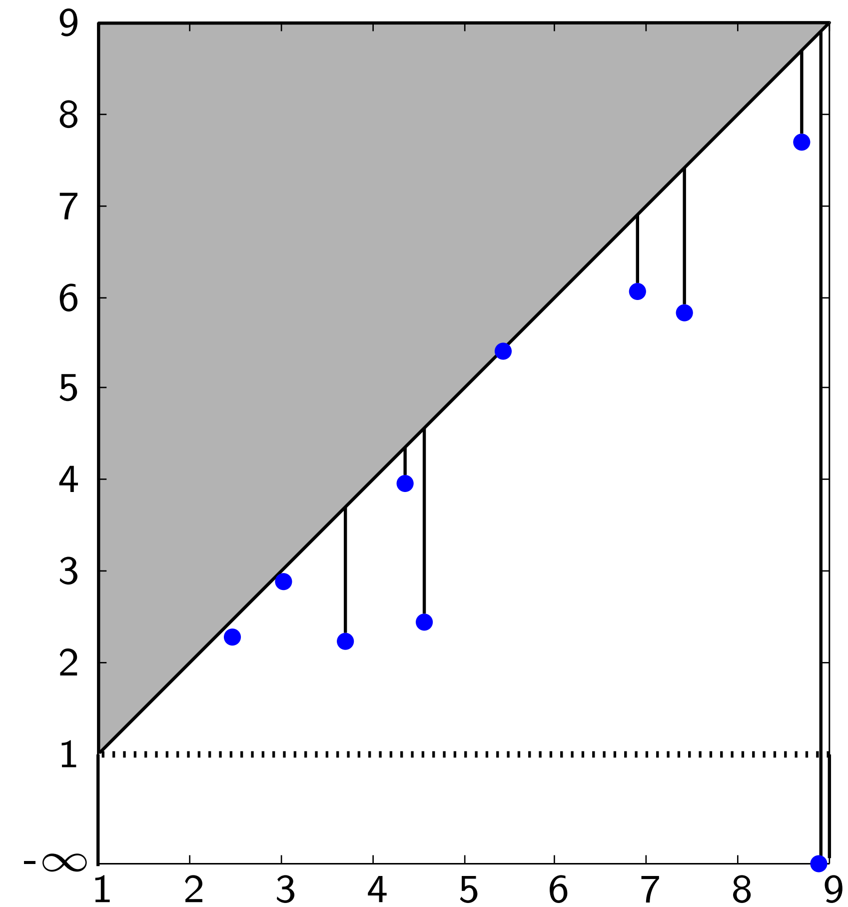
Application to Clustering

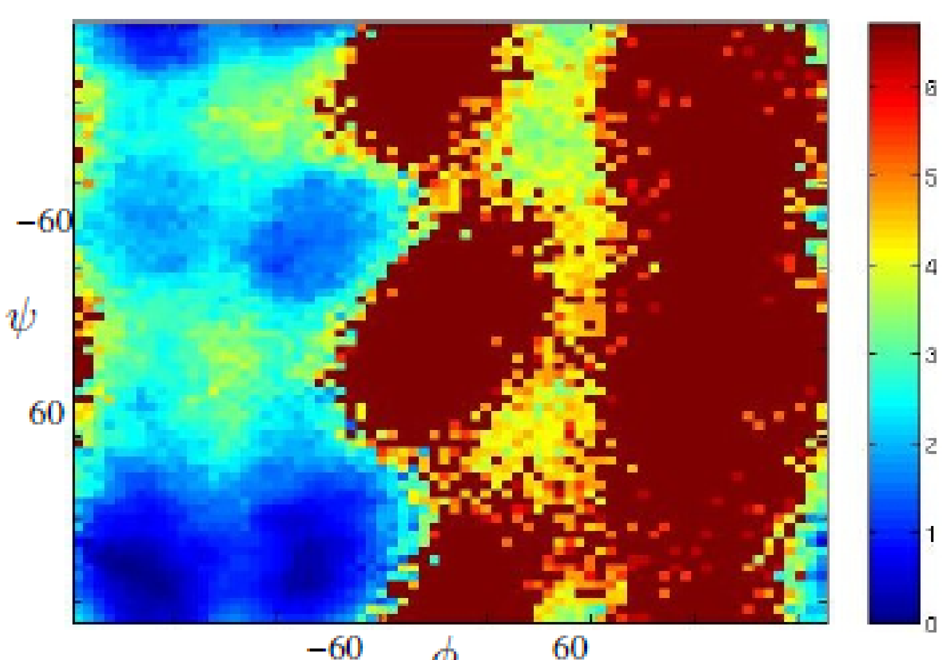
Back to the Alanine-dipeptide dataset:

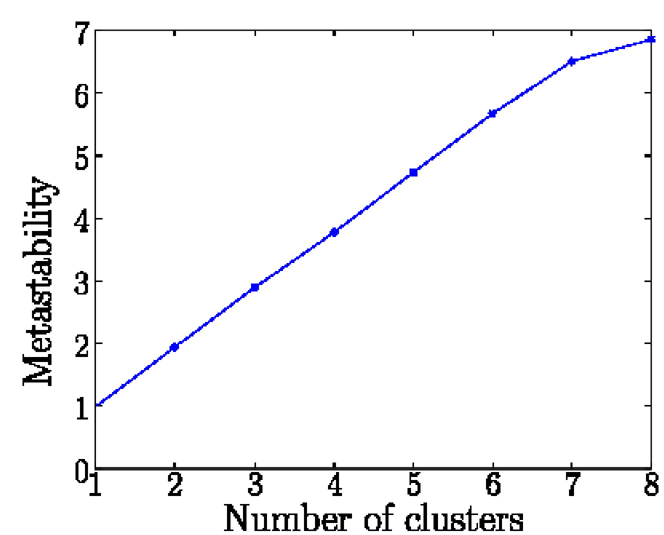
- clustering performed in R^{21}; persistence diagram plotted on a log/log scale











Interlude: Proof of Stability

$\exists\ \{\phi_\alpha : M_\alpha \to N_{\alpha+2c\delta}\}_{\alpha\in\mathbb{R}}$ and $\{\psi_\alpha : N_\alpha \to M_{\alpha+2c\delta}\}_{\alpha\in\mathbb{R}}$

s.t. the following diagrams commute $\forall \alpha \leq \beta$:

$$\begin{tikzcd}
M_{\alpha-2c\delta} \arrow[rrr] \arrow[dr, "\phi_{\alpha-2c\delta}"] & & & M_{\beta+2c\delta} \\
& N_\alpha \arrow[r] & N_\beta \arrow[ur, "\psi_\beta"] &
\end{tikzcd}$$

$$\begin{tikzcd}
& M_\alpha \arrow[r] & M_\beta \arrow[dr, "\phi_\beta"] & \\
N_{\alpha-2c\delta} \arrow[rrr] \arrow[ur, "\psi_{\alpha-2c\delta}"] & & & N_{\beta+2c\delta}
\end{tikzcd}$$

$$\begin{tikzcd}
& M_{\alpha+2c\delta} \arrow[r] & M_{\beta+2c\delta} \\
N_\alpha \arrow[ur, "\psi_\alpha"] \arrow[r] & N_\beta \arrow[ur, "\psi_\beta"] &
\end{tikzcd}$$

$$\begin{tikzcd}
M_\alpha \arrow[r] \arrow[dr, "\phi_\alpha"] & M_\beta \arrow[dr, "\phi_\beta"] & \\
& N_{\alpha+2c\delta} \arrow[r] & N_{\beta+2c\delta}
\end{tikzcd}$$

Interlude: Proof of Stability

$\exists$ $\{\phi_\alpha : M_\alpha \to N_{\alpha+2c\delta}\}_{\alpha\in\mathbb{R}}$ and $\{\psi_\alpha : N_\alpha \to M_{\alpha+2c\delta}\}_{\alpha\in\mathbb{R}}$

s.t. the following diagrams commute $\forall \alpha \leq \beta$:

$$\begin{tikzcd}
M_{\alpha-2c\delta} \arrow[rrr] \arrow[dr, "\phi_{\alpha-2c\delta}"] & & & M_{\beta+2c\delta} \\
 & N_\alpha \arrow[r] & N_\beta \arrow[ur, "\psi_\beta"] &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & M_{\alpha+2c\delta} \arrow[r] & M_{\beta+2c\delta} \\
N_\alpha \arrow[ur, "\psi_\alpha"] \arrow[r] & N_\beta \arrow[ur, "\psi_\beta"] &
\end{tikzcd}$$

$$\begin{tikzcd}
 & M_\alpha \arrow[r] & M_\beta \arrow[dr, "\phi_\beta"] & \\
N_{\alpha-2c\delta} \arrow[ur, "\psi_{\alpha-2c\delta}"] \arrow[rrr] & & & N_{\beta+2c\delta}
\end{tikzcd}
\qquad
\begin{tikzcd}
M_\alpha \arrow[r] \arrow[dr, "\phi_\alpha"'] & M_\beta \arrow[dr, "\phi_\beta"] & \\
 & N_{\alpha+2c\delta} \arrow[r] & N_{\beta+2c\delta}
\end{tikzcd}$$

Fix $\alpha_0 = 0$ and consider the following commutative diagram:

$$\begin{tikzcd}
\cdots \arrow[r] & M_0 \arrow[rr] \arrow[dr, "\phi_0"] & & M_{4c\delta} \arrow[rr] \arrow[dr, "\phi_{4c\delta}"] & & M_{8c\delta} \arrow[r] \arrow[dr, "\phi_{8c\delta}"] & \cdots \\
\cdots \arrow[ur, "\psi_{-2c\delta}"] \arrow[rr] & & N_{2c\delta} \arrow[ur, "\psi_{2c\delta}"] \arrow[rr] & & N_{6c\delta} \arrow[ur, "\psi_{6c\delta}"] \arrow[rr] & & \cdots
\end{tikzcd}$$

Interlude: Proof of Stability

- $\{{\color{blue}M_{4kc\delta}}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{M_\alpha\}_{\alpha\in\mathbb{R}}$

$$
\begin{tikzcd}
 & {\color{blue}M_0} \arrow[rr] \arrow[dr, "\phi_0"] & & {\color{blue}M_{4c\delta}} \arrow[rr] \arrow[dr, "\phi_{4c\delta}"] & & {\color{blue}M_{8c\delta}} \arrow[r] \arrow[dr, "\phi_{8c\delta}"] & \cdots \\
\cdots \arrow[ur, "\psi_{-2c\delta}"] \arrow[rr] & & {\color{red}N_{2c\delta}} \arrow[ur, "\psi_{2c\delta}"] \arrow[rr] & & {\color{red}N_{6c\delta}} \arrow[ur, "\psi_{6c\delta}"] \arrow[rr] & & \cdots
\end{tikzcd}
$$

Interlude: Proof of Stability

- $\{\color{blue}{M_{4kc\delta}}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{M_\alpha\}_{\alpha\in\mathbb{R}}$
- $\{\color{red}{N_{(4k+2)c\delta}}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{N_\alpha\}_{\alpha\in\mathbb{R}}$

$$
\begin{tikzcd}
\cdots \arrow[r] & \color{blue}{M_0} \arrow[rr] \arrow[dr, "\phi_0"] & & \color{blue}{M_{4c\delta}} \arrow[rr] \arrow[dr, "\phi_{4c\delta}"] & & \color{blue}{M_{8c\delta}} \arrow[r] \arrow[dr, "\phi_{8c\delta}"] & \cdots \\
\cdots \arrow[ur, "\psi_{-2c\delta}"] \arrow[rr] & & \color{red}{N_{2c\delta}} \arrow[ur, "\psi_{2c\delta}"] \arrow[rr] & & \color{red}{N_{6c\delta}} \arrow[ur, "\psi_{6c\delta}"] \arrow[rr] & & \cdots
\end{tikzcd}
$$

Interlude: Proof of Stability

- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{M_\alpha\}_{\alpha\in\mathbb{R}}$

- $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{N_\alpha\}_{\alpha\in\mathbb{R}}$

- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ and $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ are $4c\delta$-*discretizations* of

$$\cdots \to M_0 \to N_{2c\delta} \to M_{4c\delta} \to N_{6c\delta} \to M_{8c\delta} \to \cdots$$

$$\begin{tikzcd}
\cdots \arrow[r] & M_0 \arrow[rr] \arrow[dr, "\phi_0"] & & M_{4c\delta} \arrow[rr] \arrow[dr, "\phi_{4c\delta}"] & & M_{8c\delta} \arrow[r] \arrow[dr, "\phi_{8c\delta}"] & \cdots \\
\cdots \arrow[ur, "\psi_{-2c\delta}"] \arrow[rr] & & N_{2c\delta} \arrow[ur, "\psi_{2c\delta}"] \arrow[rr] & & N_{6c\delta} \arrow[ur, "\psi_{6c\delta}"] \arrow[rr] & & \cdots
\end{tikzcd}$$

Interlude: Proof of Stability

- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{M_\alpha\}_{\alpha\in\mathbb{R}}$

- $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{N_\alpha\}_{\alpha\in\mathbb{R}}$

- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ and $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ are $4c\delta$-*discretizations* of

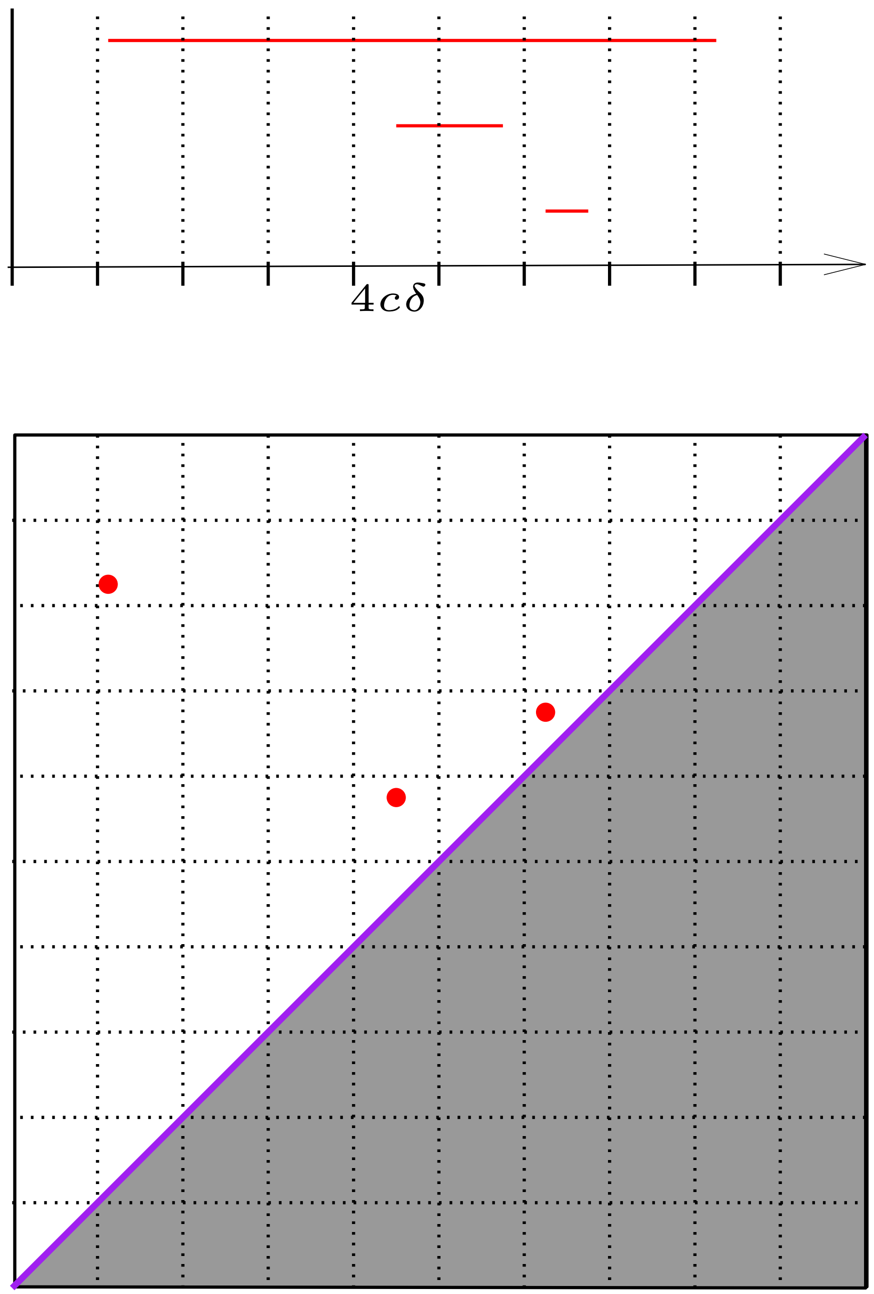
$$\cdots \to M_0 \to N_{2c\delta} \to M_{4c\delta} \to N_{6c\delta} \to M_{8c\delta} \to \cdots$$

$\to$ **goal:** relate diagrams of persistence module and of its $4c\delta$-discretizations

$$\begin{tikzcd}
\cdots \arrow[r] & M_0 \arrow[rr] \arrow[dr, "\phi_0"] & & M_{4c\delta} \arrow[rr] \arrow[dr, "\phi_{4c\delta}"] & & M_{8c\delta} \arrow[r] \arrow[dr, "\phi_{8c\delta}"] & \cdots \\
\cdots \arrow[rr] \arrow[ur, "\psi_{-2c\delta}"] & & N_{2c\delta} \arrow[rr] \arrow[ur, "\psi_{2c\delta}"] & & N_{6c\delta} \arrow[rr] \arrow[ur, "\psi_{6c\delta}"] & & \cdots
\end{tikzcd}$$

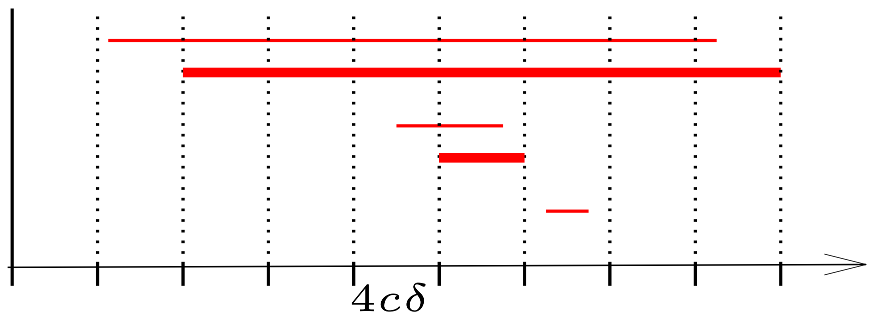
Interlude: Proof of Stability

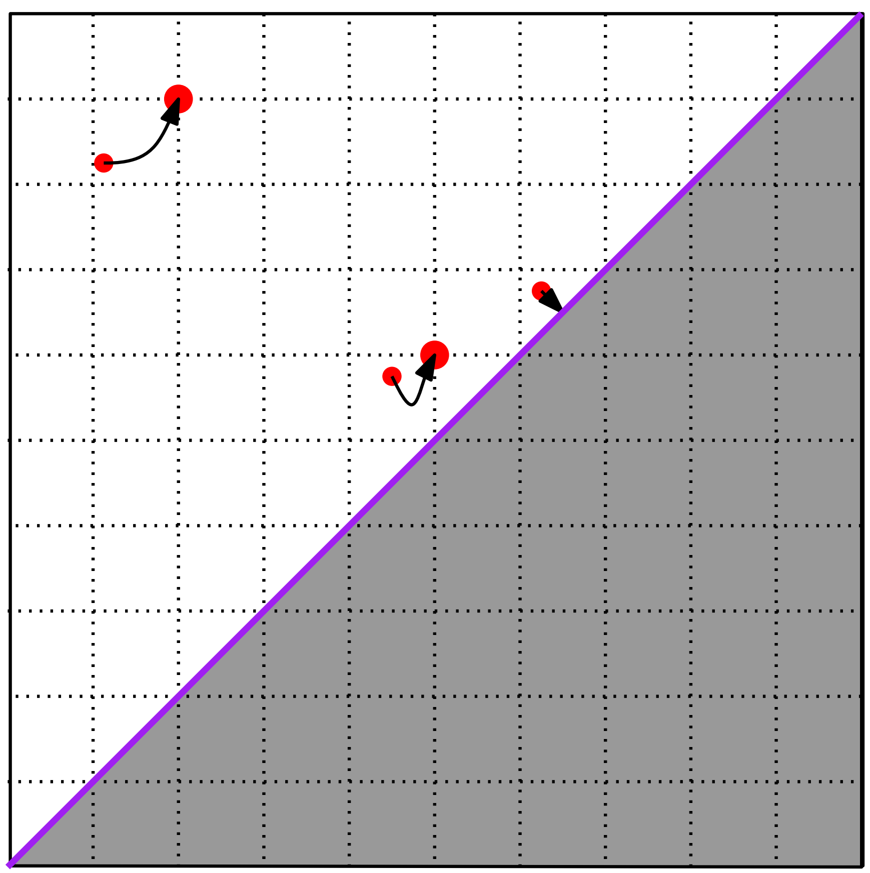
→ Effect of the discetization of a module on its persistence diagram:



Interlude: Proof of Stability

→ Effect of the discetization of a module on its persistence diagram:



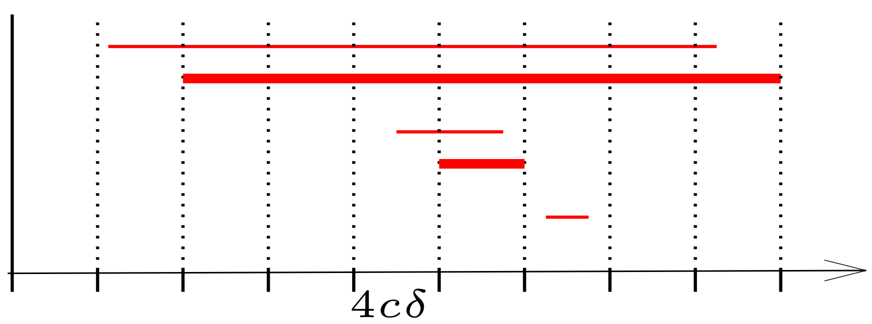


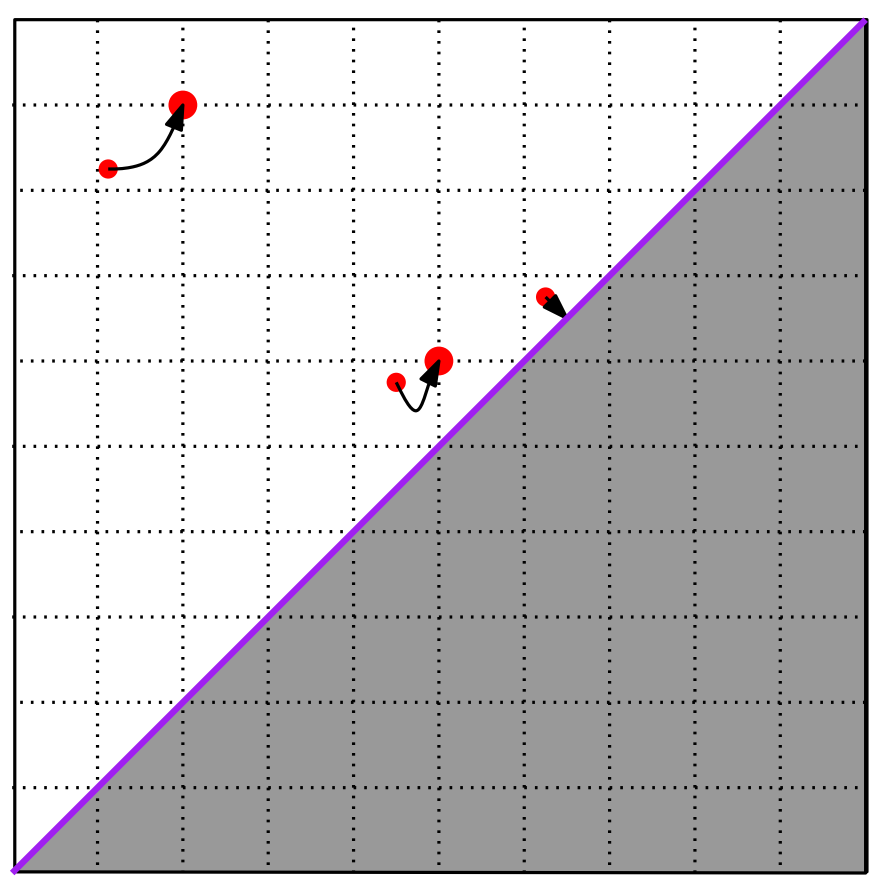
Pixelization map: $\forall \alpha \leq \beta$,

$$\pi_{4c\delta}(\alpha,\beta) = \begin{cases} \left(\lceil \frac{\alpha}{4c\delta} \rceil 4c\delta, \lceil \frac{\beta}{4c\delta} \rceil 4c\delta\right) \text{ if } \lceil \frac{\beta}{4c\delta} \rceil > \lceil \frac{\alpha}{4c\delta} \rceil \\ \left(\frac{\alpha+\beta}{2}, \frac{\alpha+\beta}{2}\right) \text{ if } \lceil \frac{\beta}{4c\delta} \rceil = \lceil \frac{\alpha}{4c\delta} \rceil \end{cases}$$

Interlude: Proof of Stability

$\longrightarrow$ Effect of the discetization of a module on its persistence diagram:





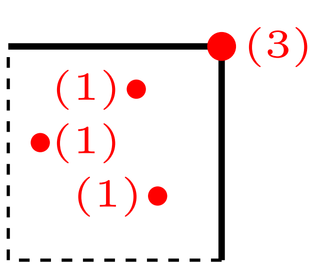
Pixelization map: $\forall \alpha \leq \beta$,

$$\pi_{4c\delta}(\alpha,\beta) = \begin{cases} \left(\lceil \frac{\alpha}{4c\delta} \rceil 4c\delta, \lceil \frac{\beta}{4c\delta} \rceil 4c\delta\right) \text{ if } \lceil \frac{\beta}{4c\delta} \rceil > \lceil \frac{\alpha}{4c\delta} \rceil \\ \left(\frac{\alpha+\beta}{2}, \frac{\alpha+\beta}{2}\right) \text{ if } \lceil \frac{\beta}{4c\delta} \rceil = \lceil \frac{\alpha}{4c\delta} \rceil \end{cases}$$

Theorem: $\pi_{4c\delta}$ induces a multi-bijection

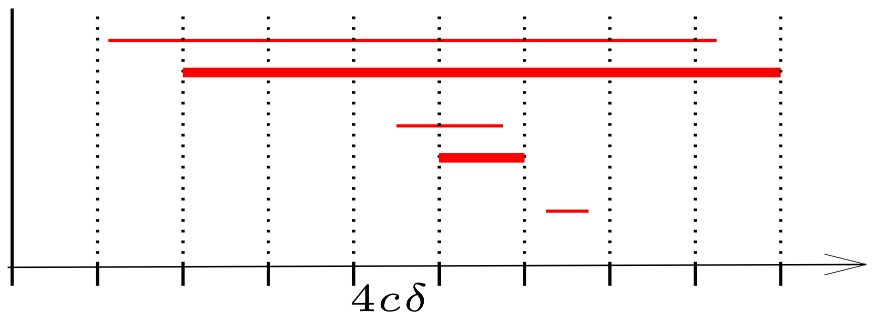
$\longrightarrow$ proof: show that the multiplicities of both diagrams are the same inside each half-open grid cell that does not intersect the diagonal.

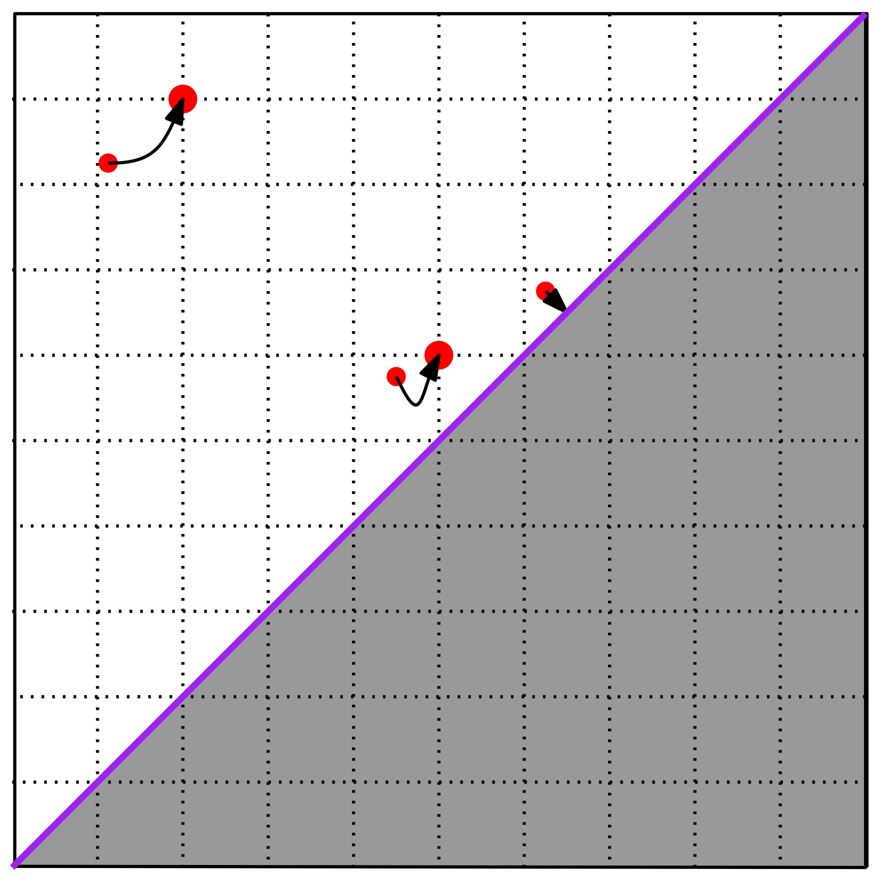
The case of diagonal cells is trivial.



Interlude: Proof of Stability

$\longrightarrow$ Effect of the discetization of a module on its persistence diagram:





Pixelization map: $\forall \alpha \leq \beta$,

$$\pi_{4c\delta}(\alpha,\beta) = \begin{cases} (\lceil \frac{\alpha}{4c\delta} \rceil 4c\delta, \lceil \frac{\beta}{4c\delta} \rceil 4c\delta) \text{ if } \lceil \frac{\beta}{4c\delta} \rceil > \lceil \frac{\alpha}{4c\delta} \rceil \\ (\frac{\alpha+\beta}{2}, \frac{\alpha+\beta}{2}) \text{ if } \lceil \frac{\beta}{4c\delta} \rceil = \lceil \frac{\alpha}{4c\delta} \rceil \end{cases}$$

Theorem: $\pi_{4c\delta}$ induces a multi-bijection

$\Rightarrow$ the bottleneck distance is at most $4c\delta$

Interlude: Proof of Stability

- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{M_\alpha\}_{\alpha\in\mathbb{R}}$
- $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{N_\alpha\}_{\alpha\in\mathbb{R}}$
- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ and $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ are $4c\delta$-*discretizations* of

$$\cdots \to M_0 \to N_{2c\delta} \to M_{4c\delta} \to N_{6c\delta} \to M_{8c\delta} \to \cdots$$

$\to$ theorem $+$ triangle inequality $\Rightarrow \mathrm{d}_{\mathrm{B}}^{\infty}(\mathrm{Dgm}\,\{M_\alpha\}_\alpha, \mathrm{Dgm}\,\{N_\alpha\}_\alpha) \leq 16c\delta$

$$\begin{tikzcd}
\cdots \arrow[r] & M_0 \arrow[rr] \arrow[dr, "\phi_0"] & & M_{4c\delta} \arrow[rr] \arrow[dr, "\phi_{4c\delta}"] & & M_{8c\delta} \arrow[r] \arrow[dr, "\phi_{8c\delta}"] & \cdots \\
\cdots \arrow[rr] \arrow[ur, "\psi_{-2c\delta}"] & & N_{2c\delta} \arrow[rr] \arrow[ur, "\psi_{2c\delta}"] & & N_{6c\delta} \arrow[rr] \arrow[ur, "\psi_{6c\delta}"] & & \cdots
\end{tikzcd}$$

Interlude: Proof of Stability

- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{M_\alpha\}_{\alpha\in\mathbb{R}}$

- $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ is a $4c\delta$-*discretization* of $\{N_\alpha\}_{\alpha\in\mathbb{R}}$

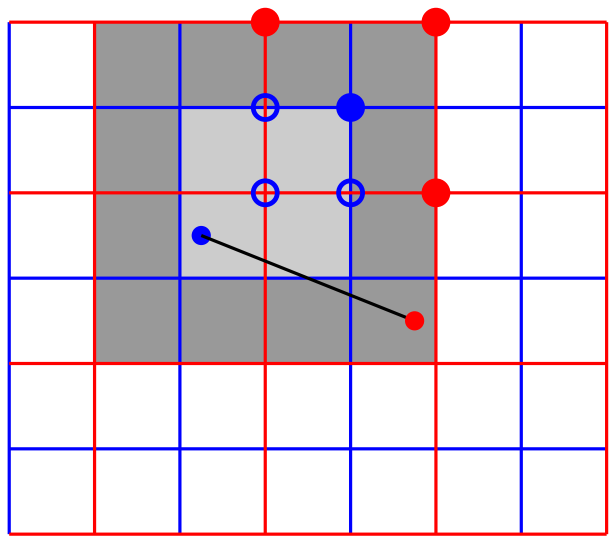
- $\{M_{4kc\delta}\}_{k\in\mathbb{Z}}$ and $\{N_{(4k+2)c\delta}\}_{k\in\mathbb{Z}}$ are $4c\delta$-*discretizations* of

  $$\cdots \to M_0 \to N_{2c\delta} \to M_{4c\delta} \to N_{6c\delta} \to M_{8c\delta} \to \cdots$$

$\to$ theorem + triangle inequality $\Rightarrow \mathrm{d}_{\mathrm{B}}^{\infty}(\mathrm{Dgm}\,\{M_\alpha\}_\alpha, \mathrm{Dgm}\,\{N_\alpha\}_\alpha) \le 16c\delta$

Improvement:

$$\mathrm{d}_{\mathrm{B}}^{\infty}(\mathrm{Dgm}\,\{M_\alpha\}_\alpha, \mathrm{Dgm}\,\{N_\alpha\}_\alpha) \le 6c\delta$$



Some References

[1] Chazal, Cohen-Steiner, Glisse, Guibas, O., *Proximity of persistence modules and their diagrams*, Proc. SoCG 2009.

[2] Chazal, Guibas, O., Skraba, *Analysis of scalar fields over point cloud data*, Proc. SODA 2009.

[3] Chazal, Guibas, O., Skraba, "Persistence-Based Clustering in Riemannian Manifolds", INRIA Research report 6968, June 2009.

$\rightarrow$ code integrated to the CULT library (many thanks to P. Skraba)
`https://gforge.inria.fr/projects/cult/`