Persistent Intrinsic Volumes

Persistent homology as a tool for geometric inference

Antoine Commaret
David Cohen-Steiner

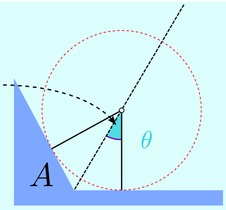
Inria — Université Côte d'Azur

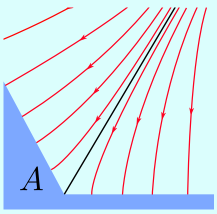
Problem. Recovering the boundary area of a set Y from an approximating set X as a function of $\varepsilon = d_H(X, Y)$.

Some notations. d_A denotes the distance to any $A \subset \mathbb{R}^d$.

- $A^t := \{x \in \mathbb{R}^d \mid d_A(x) \leq t\}$ is the t-offset of A.
- $H_i(A)$ is the i-th homology vector space over an arbitrary field.
- $\dim H_i(A)$ is its i-th Betti number.

The **μ-reach** of a set A is the largest t such that for any $x \in A^t \setminus A$, the cosine of the half-angle between two closest points of x in X is less than μ.





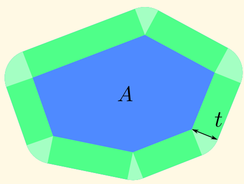
Theorem. When $t < \mathrm{reach}_\mu(A)$, for every $\varepsilon > 0$, there exists a flow whose trajectories, parametrized by the arc-length, make d_A decrease at speed $\geq \mu - \varepsilon$ on $A^t \setminus A$.

Smooth sets, polyhedra and most stratified sets have positive μ-reach.

Intrinsic volumes $V_0(A), \ldots V_d(A)$ are quantities defined for most subsets A of $\mathbb{R}^d$, and are related to their curvatures. For submanifolds, they coincide with the *Lipschitz-Killing curvatures.* For convex sets, they are defined by the volume of offsets.

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

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



More generally, they can be obtained via the **principal kinematic formula**.

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

Steiner's polynomial of A

Example. Up to scaling constants,

- $V_{d-1}(A)$ is the boundary area $\mathcal{H}^{d-1}(\partial A)$.
- $V_1(A)$ is the mean curvature of A.
- $V_0(A)$ is the Euler characteristic $\chi(A)$.

The **persistent Steiner polynomial** is defined as

$$Q_{X,\varepsilon}(t) := \int_{\mathbb{R}^d} \chi(\mathrm{dgm}(X, \varepsilon, x)_t)\,\mathrm{d}x$$

Final theorem. Recovering surrogates coefficient from $Q_{X,\varepsilon}$ yields quantities $V_i(X, \varepsilon, R)$ such that

$$\left|V_i(X,\varepsilon,R) - V_i(Y^{2\varepsilon})\right| \leq \frac{P(i,d,R)M_R(Y^{2\varepsilon})\varepsilon}{\mu}$$

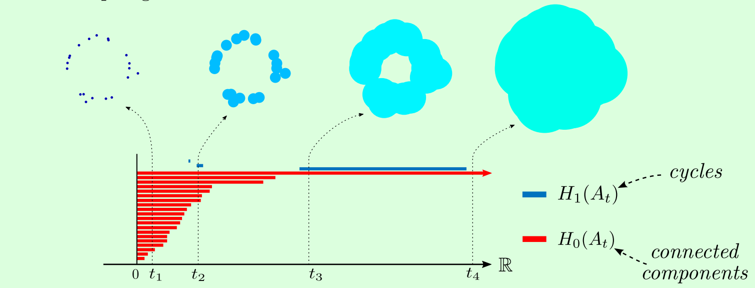
fraction in i, d, R

When Y has bounded total curvatures, we further show that

$$|V_i(X,\varepsilon,R) - V_i(Y)| \leq \left(M_R(Y^{2\varepsilon}) + M_R(Y)\right)\frac{P(i,d,R)\varepsilon}{\mu}$$

A **filtration** is a family $(A_t)_{t\in\mathbb{R}}$ non-decreasing for the inclusion.

A **persistence diagram** associates to a filtration the values of birth and death of topological features.



Diagrams are made rigorous by their algebraic counterparts, called **persistent homology modules**. In dimension i, a bar **is born** (resp. **dies**) at time t when the dimension of $H_i(A_t)$ **increases** (resp. **decreases**) at t.

$$\begin{tikzcd} {} \arrow[r, hook] & A_t \arrow[r, hook, "\phi"] \arrow[d, phantom, "\Downarrow"] & A_s \arrow[r, hook] & {} \\ {} \arrow[r] & H_i(A_t) \arrow[r, "\phi_*"'] & H_i(A_s) \arrow[r] & {} \end{tikzcd}$$

inclusion-induced linear map between vector spaces

If $(A_t), (B_t)$ are two filtrations such that for every t, $A_t \subset B_t$, the **image persistence module** $dgm(A, B)$ is given by the bars of the families $\iota_*(H_i(A_t))$.

$$\begin{tikzcd} {} \arrow[r] & \iota_*(H_i(A_t)) \arrow[r] \arrow[d] & \iota_*(H_i(A_s)) \arrow[r] \arrow[d] & {} \\ {} \arrow[r] & H_i(B_t) \arrow[r] & H_i(B_s) \arrow[r] & {} \\ {} \arrow[r] & H_i(A_t) \arrow[r] \arrow[u, "\iota_*"] \arrow[uu, dashed, bend right] & H_i(A_s) \arrow[r] \arrow[u, "\iota_*"'] \arrow[uu, dashed, bend right] & {} \end{tikzcd}$$

Since $X^\varepsilon \subset Y^{2\varepsilon} \subset X^{3\varepsilon}$, for any $x \in \mathbb{R}^d$ we put $\mathrm{dgm}(X, \varepsilon, x)$ the persistence image diagram induced by the two filtrations $(X^\varepsilon \cap B(x,t))_{t\in\mathbb{R}}$ and $(X^{3\varepsilon} \cap B(x,t))_{t\in\mathbb{R}}$, and $\mathrm{dgm}(Y^{2\varepsilon}, x)$ for $(Y^{2\varepsilon} \cap B(x,t))_{t\in\mathbb{R}}$

$$\begin{tikzcd} {} \arrow[r] & H_i(X^{3\varepsilon} \cap B(x,t)) \arrow[r] & H_i(X^{3\varepsilon} \cap B(x,s)) \arrow[r] & {} \\ {} \arrow[r] & H_i(Y^{2\varepsilon} \cap B(x,t)) \arrow[r] \arrow[u] & H_i(Y^{2\varepsilon} \cap B(x,s)) \arrow[r] \arrow[u] & {} \\ {} \arrow[r] & H_i(X^{\varepsilon} \cap B(x,t)) \arrow[r] \arrow[u] & H_i(X^{\varepsilon} \cap B(x,s)) \arrow[r] \arrow[u] & {} \end{tikzcd}$$

χ-averaging Lemma : Let $\mu > 0$ such that $d_H(X, Y) \leq \varepsilon \leq \mathrm{reach}_\mu(X)$. Then

$$\int_0^R \left|\chi(\mathrm{dgm}(X,\varepsilon,x)_t) - \chi(\mathrm{dgm}(Y^{2\varepsilon},x)_t)\right| \mathrm{d}t \leq 2\frac{\varepsilon}{\mu} N_0^R(\mathrm{dgm}(Y^{2\varepsilon}, x))$$

Alternating sum of the Betti numbers at filtration value t

Number of bars of the diagram inside $[0, R]$

Theorem. Let $\mu > 0$ be such that $d_H(X, Y) \leq \varepsilon \leq \frac{1}{4}\mathrm{reach}_\mu(X)$. Then,

$$\int_0^R |Q_{X,\varepsilon}(t) - Q_{Y^{2\varepsilon}}(t)|\,\mathrm{d}t \leq \frac{4\varepsilon}{\mu}\int_{\mathbb{R}^d} N_0^R(\mathrm{dgm}(Y^{2\varepsilon}, x))$$

Morse theory Lemma. Counting the critical points of $(d_x)_{|Y}$ yields:

$$\int_{\mathbb{R}^d} N_0^R(\mathrm{dgm}(Y^{2\varepsilon}, x)) \leq M_R(Y^{2\varepsilon})$$

depends on the total curvatures of $Y^{2\varepsilon}$.