Critical points and sampling theory of generic submanifolds

Charles Arnal[1], David Cohen-Steiner[1] & Vincent Divol[2]

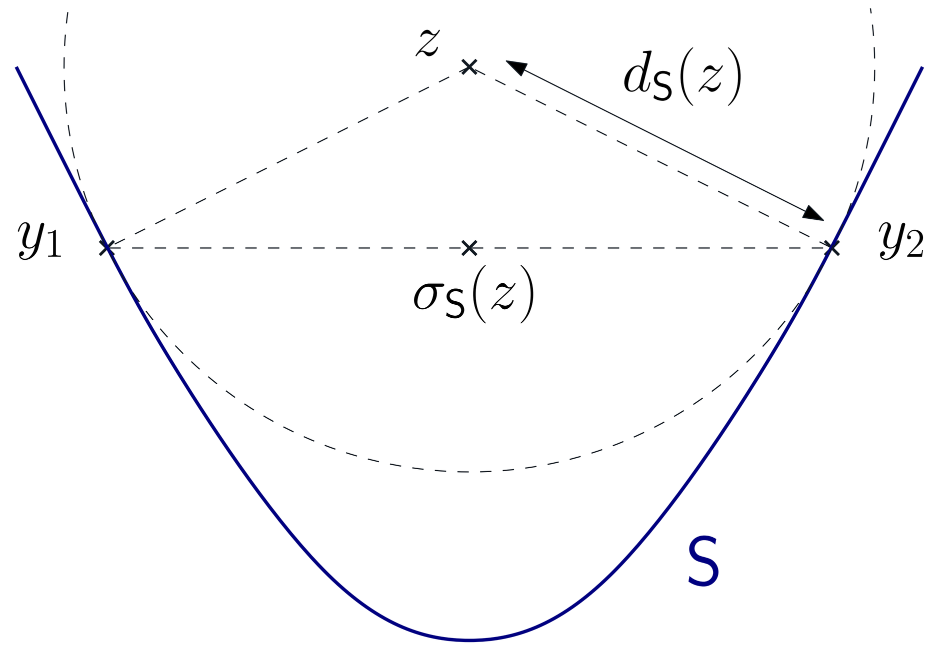
Datashape, Inria[1] & CREST, ENSAE[2]

High-Level Summary

We are interested in C^2 compact submanifolds $\mathsf{M} \subset \mathbb{R}^d$ and random finite samplings $\mathsf{A} \subset \mathsf{M}$, in the distance functions d_M and d_A and in their critical points and their persistence diagrams $\mathrm{dgm}_i(\mathsf{M})$ and $\mathrm{dgm}_i(\mathsf{A})$. These can be very irregular in general, but we show that they are well-behaved when M is generic.

Definitions

Let $\Pi_\mathsf{S}(z)$ denote the set $\{x \in \mathsf{S} : d(z,x) = d_\mathsf{S}(z)\}$ of *projections* of z on a closed set $\mathsf{S} \subset \mathbb{R}^d$, and let $\sigma_\mathsf{S}(z)$ denote the projection of z onto the convex hull $\mathrm{Conv}(\Pi_\mathsf{S}(z))$. The *generalized gradient* of d_S is $\nabla d_\mathsf{S}(z) = \frac{z - \sigma_\mathsf{S}(z)}{d_\mathsf{S}(z)}$ if $z \notin \mathsf{S}$, and $\nabla d_\mathsf{S}(z) = 0$ if $z \in \mathsf{S}$. $\mathrm{Crit}(\mathsf{S}) := \{z \in \mathbb{R}^d \backslash \mathsf{S} : \nabla d_\mathsf{S}(z) = 0\}$ is the set of *critical points* of S.



Genericity Theorem

For a generic C^2 submanifold M, i.e. for a dense and open set of embeddings of any abstract manifold with respect to the C^2 topology, $\mathrm{Crit}(\mathsf{M})$ and $\cup_{z \in \mathrm{Crit}(\mathsf{M})} \Pi_\mathsf{M}(z)$ are well-behaved:

(P1) (Non-degeneracy of the projections) For every $z_0 \in \mathrm{Crit}(\mathsf{M})$, the projections $\pi_\mathsf{M}(z_0)$ are the vertices of a non-degenerate simplex of $\mathbb{R}^d$ and z_0 belongs to the relative interior of $\mathrm{Conv}(\pi_\mathsf{M}(z_0))$.

(P2) (Finiteness of $\mathrm{Crit}(\mathsf{M})$) The set $\mathrm{Crit}(\mathsf{M})$ is finite.

(P3) (Tangency condition) For every $z_0 \in \mathrm{Crit}(\mathsf{M})$ and every $x_0 \in \pi_\mathsf{M}(z_0)$, the sphere $S(z_0, d_\mathsf{M}(z_0))$ is non-osculating M at x_0 (i.e. it is not tangent of order 2).

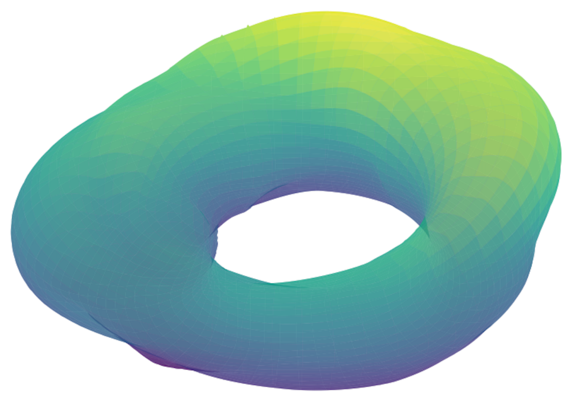
(P4) (Morse-like behaviour) d_M is a topological Morse function, and as such satisfies the Isotopy Lemma and the Handle Attachment Lemma.

$\rightarrow$ As a result, the Čech persistence diagram of degree i of M, $\mathrm{dgm}_i(\mathsf{M})$, is finite and can be entirely recovered from $\mathrm{Crit}(\mathsf{M})$: each point $z_0 \in \mathrm{Crit}(\mathsf{M})$ of critical index i contributes to the birth of an i-cycle or to the death of an $(i-1)$-cycle.

Stability Theorem

When M is generic, $\mathrm{Crit}(\mathsf{M})$ and $\cup_{z \in \mathrm{Crit}(\mathsf{M})} \Pi_\mathsf{M}(z)$ are also stable, in the sense that for any $\varepsilon > 0$, any other embedding M' close enough to M for the C^2 topology is such that

- M' satisfies (P1-4).
- There is a bijection $\Psi : \mathrm{Crit}(\mathsf{M}) \to \mathrm{Crit}(\mathsf{M}')$.
- Let $z_0 \in \mathrm{Crit}(\mathsf{M})$; then $\Psi(z_0) \in B(z_0, \varepsilon)$.
- Let $z_0 \in \mathrm{Crit}(\mathsf{M})$ with $\pi_\mathsf{M}(z_0) = \{x_1, \ldots, x_s\}$. Then $\pi_{\mathsf{M}'}(\Psi(z_0))$ is included in the union of balls of radius ε centered at the x_j, with exactly one point in each ball.



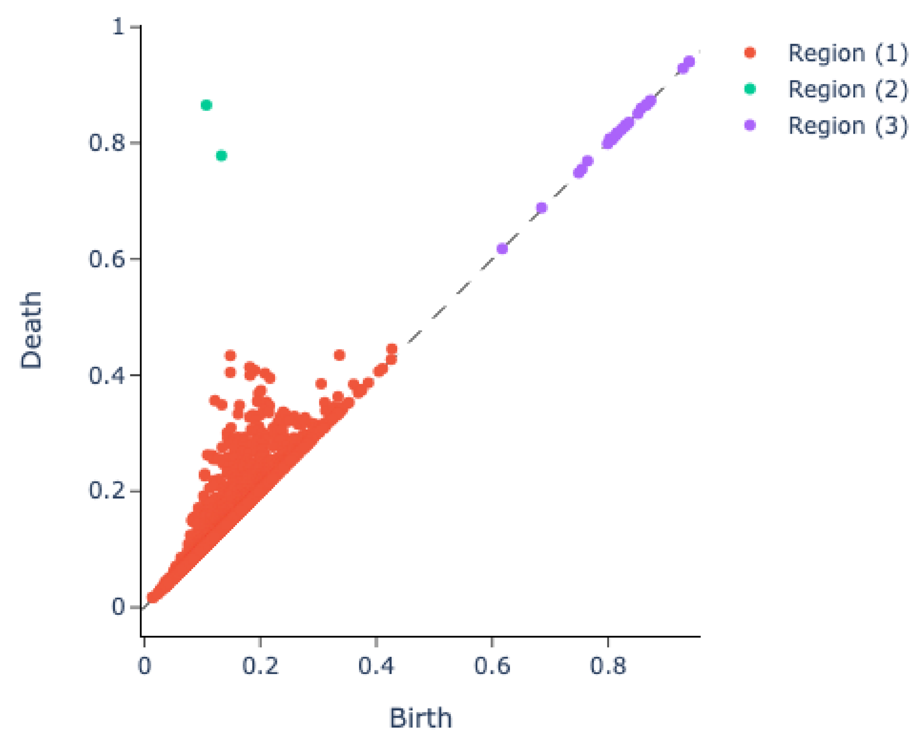


Figure: (Left) A generic torus. (Right) The (Čech) persistence diagram $\mathrm{dgm}_1(\mathsf{A}_n)$ of a set of $n = 10^4$ points sampled on a generic torus with some

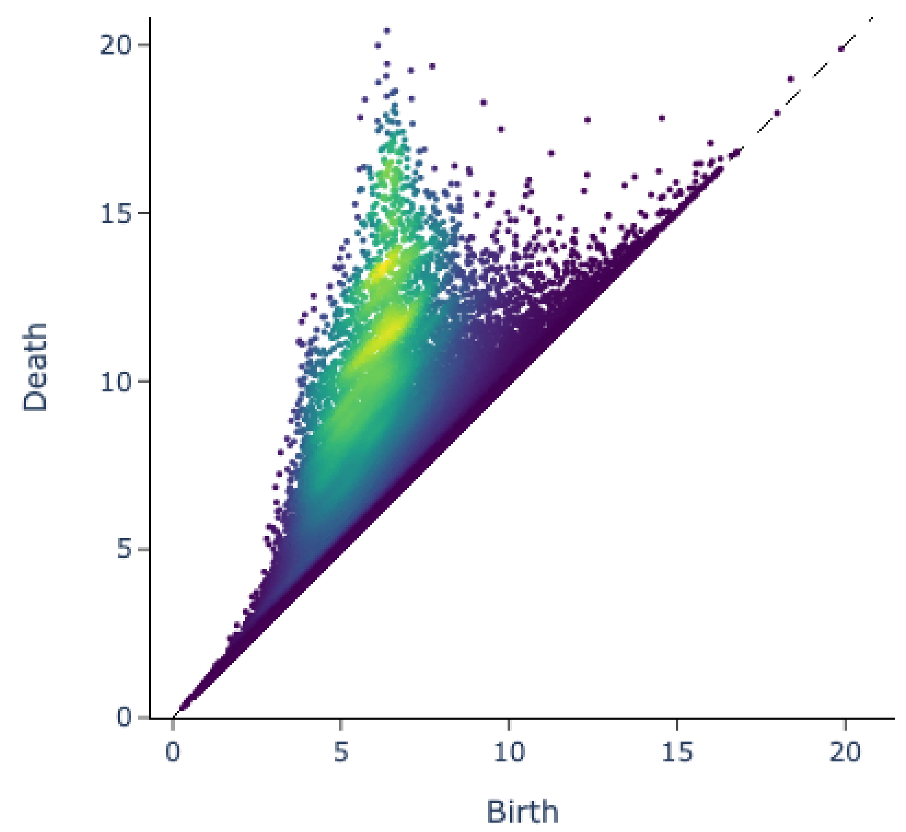


Figure: Zoom on Region 1: The limit measure $\mu_{P,1}$.

Persistence diagrams of subsamplings

Let $\mathsf{A} \subset \mathsf{M}$ be a finite sampling s.t. $d_H(\mathsf{A}, \mathsf{M}) \leq \varepsilon$. $\mathrm{dgm}_i(\mathsf{A})$ can be partitioned into three regions:

$$\mathrm{dgm}_i(\mathsf{A}) = \Big(\mathrm{dgm}_i(\mathsf{A}) \cap \{u_1, u_2 \leq \varepsilon + \frac{\varepsilon^2}{\tau(\mathsf{M})}\}\Big) \quad \Big\}\mathrm{dgm}_i^{(1)}(\mathsf{A}_n)$$
$$\sqcup \Big(\mathrm{dgm}_i(\mathsf{A}) \cap \{u_1 \leq \varepsilon, u_2 \geq \tau(\mathsf{M}) - \frac{\varepsilon^2}{\tau(\mathsf{M})}\}\Big) \quad \Big\}\mathrm{dgm}_i^{(2)}(\mathsf{A}_n)$$
$$\sqcup \Big(\mathrm{dgm}_i(\mathsf{A}) \cap \{u_1, u_2 \geq \tau(\mathsf{M}) - \frac{\varepsilon^2}{\tau(\mathsf{M})}\Big) \quad \Big\}\mathrm{dgm}_i^{(3)}(\mathsf{A}_n)$$

where $\tau(\mathsf{M})$ is the reach of M.

If $\mathsf{A} = \mathsf{A}_n$ is a random sampling drawn from a distribution P with a density bounded away from zero and infinity on M, then:

- (Region 1) Let $\mu_{n,i} = \frac{1}{n}\sum_{u \in \mathrm{dgm}_i^{(1)}(\mathsf{A}_n)} \delta_{n^{1/m}u}$. Then $\mathbb{E}[\mathrm{OT}_p^p(\mu_{n,i}, \mu_{P,i})] \xrightarrow[n\to\infty]{} 0$ for some limit measure $\mu_{P,i}$, where OT_p is a generalization of the Wasserstein distance.
- (Regions 2 and 3) $\mathbb{E}[\#(\mathrm{dgm}_i^{(3)}(\mathsf{A}_n))] = O(1)$, and there exists an optimal matching $\gamma_n : \mathrm{dgm}_i(\mathsf{A}_n) \cup \partial\Omega \to \mathrm{dgm}_i(\mathsf{M}) \cup \partial\Omega$ for the bottleneck distance such that
$$\mathbb{E}[\max_{u \in \mathrm{dgm}_i^{(2)}(\mathsf{A}_n)} |u_2 - \gamma_n(u)_2|] = O(n^{-2/m})$$
$$\mathbb{E}[\max_{u \in \mathrm{dgm}_i^{(3)}(\mathsf{A}_n)} \|u - \gamma_n(u)\|_\infty] = O(n^{-2/m}).$$

Wasserstein convergence and total persistence:

As a consequence, $\mathbb{E}[\mathrm{OT}_p^p(\mathrm{dgm}_i(\mathsf{A}_n), \mathrm{dgm}_i(\mathsf{M}))] \to 0$ as $n \to \infty$ if $p > m$, and the total persistence $\mathrm{Pers}_\alpha(\mathrm{dgm}_i(\mathsf{A}_n)) = \sum_{u \in \mathrm{dgm}_i(\mathsf{A}_n)} (\frac{u_2 - u_1}{2})^\alpha$ satisfies

$$\mathrm{Pers}_\alpha(\mathrm{dgm}_i(\mathsf{A}_n)) = \mathrm{Pers}_\alpha(\mathrm{dgm}_i(\mathsf{M})) + n^{1-\frac{\alpha}{m}} C_{i,P,\mathsf{M}} + o_{L^1}(n^{1-\frac{\alpha}{m}}) + O_{L^1}\Big(\Big(\frac{\log n}{n}\Big)^{\frac{1}{m}}\Big).$$

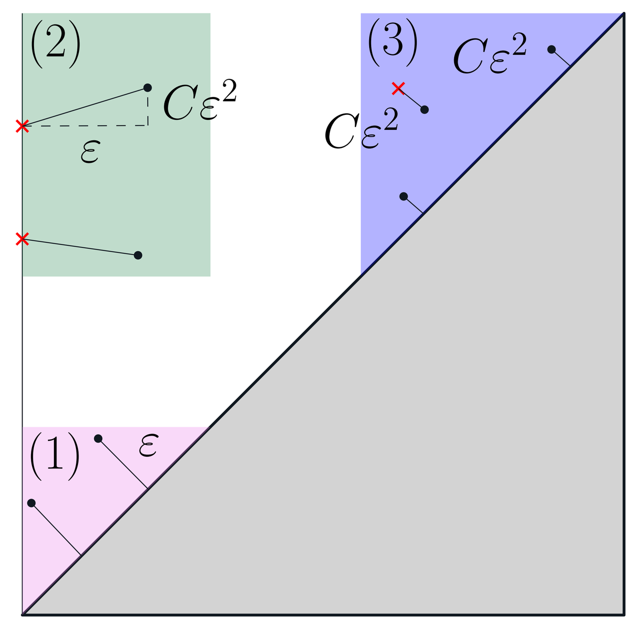


Figure: The partition of the persistence diagram into three regions.

References

- *Critical points of the distance function to a generic submanifold*, Charles Arnal, David Cohen-Steiner, Vincent Divol
- *Wasserstein convergence of Čech persistence diagrams for samplings of submanifolds*, Charles Arnal, David Cohen-Steiner, Vincent Divol