Generic families of finite metric spaces with identical or trivial 1-dimensional persistence

Philip Smith[1] · Vitaliy Kurlin[2]

Abstract

Persistent homology is a popular and useful tool for analysing finite metric spaces, revealing features that can be used to distinguish sets of unlabeled points and as input into machine learning pipelines. The famous stability theorem of persistent homology provides an upper bound for the change of persistence in the bottleneck distance under perturbations of points, but without giving a lower bound. This paper clarifies the possible limitations persistent homology may have in distinguishing finite metric spaces, which is evident for non-isometric point sets with identical persistence. We describe generic families of point sets in metric spaces that have identical or even trivial one-dimensional persistence. The results motivate stronger invariants to distinguish finite point sets up to isometry.

1 Motivations, problems, and outline of results

Topological Data Analysis (TDA) summarises geometric and topological features in unstructured data and was pioneered by Serguei Barannikov (1994), Claudia Frosini and Landi (1999), Vanessa Robins (1999), and Edelsbrunner et al. (2000). The key papers of Gunnar Carlsson (2009), Robert Ghrist (2008), and Shmuel Weinberger (2011) were followed by substantial developments of Fred Chazal et al. (2016) and others.

✉ Vitaliy Kurlin
  vitaliy.kurlin@liverpool.ac.uk

  Philip Smith
  philip.smith-2@manchester.ac.uk

[1] Henry Royce Institute, University of Manchester, Oxford Road, Manchester M13 9PL, UK

[2] Department of Computer Science, University of Liverpool, Ashton street, Liverpool L69 3BX, UK
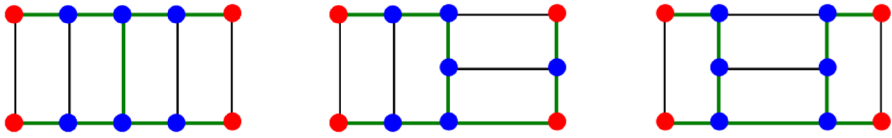
Ⓐ Springer

Fig. 1 Many non-isometric sets have the same 0D persistence and trivial 1D persistence. Theorem 4.4 extends these examples to generic families of sets by adding 'tails' at red corners (colour figure online)

The main tool of TDA (Edelsbrunner and Harer 2008) is persistent homology, which is defined below via a filtration of complexes on a *point cloud* (a finite set A of unordered points). One can also consider filtrations of sublevel sets of a scalar function.

Definition 1.1 (*A filtration of complexes* $\{C(A; \alpha)\}$) Let A be any finite set.
(a) A simplicial *complex* C on A is a finite set of subsets $\sigma \subset A$ (*simplices*) such that all subsets of $\sigma \subset A$ and hence all intersections of simplices are simplices of C.
(b) The *dimension* of a simplex σ on $k + 1$ points is k. We assume that all points of A are 0-dimensional simplices, sometimes called *vertices* of C. A 1-dimensional simplex (or *edge*) e between points $p, q \in A$ is the unordered pair denoted as $[p, q]$.
(c) An ascending *filtration* $\{C(A; \alpha)\}$ is a family of complexes on the vertex set A, paremetrised by a *scale* $\alpha \in \mathbb{R}$ so that $C(A; \alpha') \subseteq C(A; \alpha)$ for $\alpha' \leq \alpha$. ∎

Definition 1.2 (*1D persistence and barcode*) For any filtration $\{C(A; \alpha)\}$ of complexes on a cloud A in a metric space, a homology class $\gamma \in H_1(C(A; \alpha_i))$ is *born* at $\alpha_i = \text{birth}(\gamma)$ if γ is not in the full image under the induced homomorphism $H_1(C(A; \alpha)) \rightarrow H_1(C(A; \alpha_i))$ for any $\alpha < \alpha_i$. The class γ *dies* at $\alpha_j = \text{death}(\gamma) \geq \alpha_i$ when the image of γ under $H_1(C(A; \alpha_i)) \rightarrow H_1(C(A; \alpha_j))$ merges into the image under $H_1(C(A; \alpha)) \rightarrow H_1(C(A; \alpha_j))$ for some $\alpha < \alpha_i$.

Let $\alpha_1, \ldots, \alpha_m$ be all scales when a homology class is born or dies in $H_1(C(A; \alpha))$. Let μ_{ij} be the number of independent classes in $H_1(C(A; \alpha))$ that are born at α_i and die at α_j. The 1D *persistence* diagram $\text{PD}_1\{C(A; \alpha)\} \subset \mathbb{R}^2$ is the multi-set consisting of the pairs (α_i, α_j) with integer multiplicities $\mu_{ij} \geq 1$. The 1D *barcode* is the unordered multi-set of intervals $[\alpha_i, \alpha_j)$ with multiplicities μ_{ij}. ∎

The birth-death pairs from Definition 1.2 can be similarly defined for any k-dimensional homology groups H_k with $k \geq 0$ and coefficients in a field, though the coefficients in $\mathbb{Z}_2 = \{0, 1\}$ are used in practice and in this paper.

Standard filtrations of (Vietoris-Rips, Čech and Delaunay) complexes on a point cloud A are introduced in Definition 2.1. For all these filtrations in dimension 0, the homology group $H_0(C(A; \alpha))$ is generated by the single-linkage clusters of A formed by all points that can be connected through inter-point distances up to 2α. Then all homology classes of $H_0(C(A; \alpha))$ are born at $\alpha = 0$ (isolated dots) and die at α equal to half-lengths of all edges in a Minimum Spanning Tree MST(A). Figure 1 shows the edges of MST(A) in green: one edge of length 2 and eight edges of length 1 for each vertex set A.

In all examples of Fig. 1, the 1-dimensional persistence is trivial (empty due to no pairs with death > birth) because all rectangular 'dominoes' do not create cycles in the 1D homology for the standard filtrations of complexes.
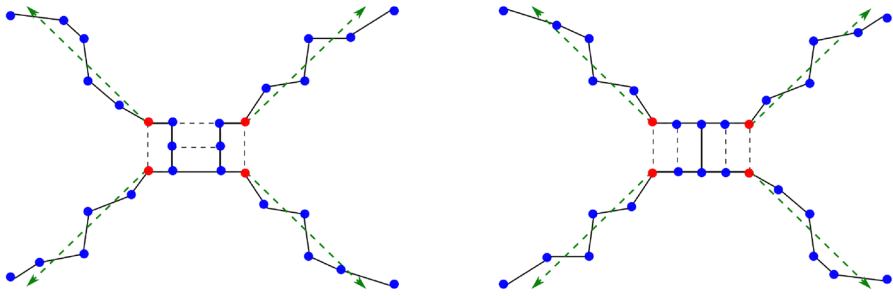
Fig. 2 The set *A* of 10 points in the centre is extended by four tails going out from red points. All such sets have trivial 1D persistence by Corollary 4.5, but all such sets in general position are not isometric to each other. The black edges form a minimum spanning tree (colour figure online)

In any dimension for filtrations based only on inter-point distances, the resulting persistence diagram is invariant under *isometry* preserving inter-point distances, not up to more general continuous deformations.

Hence most persistence-based classifications distinguish point clouds only up to isometry, which is an important equivalence due to the rigidity of many real-life structures. Figure 1 shows sets $A \subset \mathbb{R}^2$ whose points (in blue and red) form 1×2 'dominoes' that have identical persistence in dimensions 0 and 1.

To understand the strength of persistence as an isometry invariant, the following problem asks to fully describe the inverse of the persistence map.

Problem 1.3 (*Inverting persistence*) For a given filtration of complexes, find necessary and sufficient conditions for finite metric spaces to have a given persistence diagram in each dimension. In particular, describe all *1D homologically trivial* point sets that (by definition) have a trivial (empty) 1D persistence diagram. ∎

The analogue of Problem 1.3 was solved for 0-dimensional persistence of Morse-like functions on the interval (Curry 2018), see also (Catanzaro et al. 2020). Main Theorem 4.4 will show how any number of points can be added to any finite point set whilst leaving the 1-dimensional persistence unchanged, see Fig. 2 extending Fig. 1.

Corollary 4.5 describes generic families of finite metric spaces that have trivial 1-dimensional persistence for the standard filtrations of simplicial (Vietoris-Rips, Čech and Delaunay) complexes in Definition 2.1. For high-dimensional data, usually only the Vietoris-Rips complex (determined by its 1D skeleton) is computationally feasible (Bauer 2021). Problem 1.3 for point clouds and persistence in dimensions more than 1 will be discussed in future work.

In the context of Problem 1.3, the resulting families of point clouds in $\mathbb{R}^N$ form vast open subspaces (in the space of isometry classes of all point clouds in $\mathbb{R}^N$), which are mapped by to a single 1D persistence diagram. This result complements the famous stability theorem (Cohen-Steiner et al. 2005) stating that under bounded noise, the bottleneck distance between persistence diagrams of a point set and its perturbation has an upper bound depending on the magnitude of the perturbation. However, there is no lower bound, so a perturbation of a point set can result in the corresponding persistent homology remaining unchanged.

Section 4 introduces definitions and proves auxiliary lemmas needed for our main Theorem 4.4, which describes how, given a finite point set, we can add an arbitrarily large point set without affecting the one-dimensional persistent homology. Section 5 summarises large-scale experiments that reveal important information on the prevalence, or more likely lack, of significant persistent features occurring in randomly generated point clouds in many dimensions.

Since persistence is preserved under small perturbations of many point clouds, we might be interested in stronger isometry invariants discussed in Sect. 6. Indeed, many applications (Edelsbrunner et al. 2021) need to reliably distinguish point sets up to isometry or similar equivalence relations such as rigid motion or uniform scaling. A uniform scaling also scales persistence, but a more general continuous deformation of data changes persistence rather arbitrarily.

2 Edges that are important for 1D persistence

This section introduces three classes of edges (short, medium, and long) that will help build point sets with identical 1D persistence. Since persistent homology can be defined for any filtration of simplicial complexes on an abstract finite set A, the most general settings are recalled in Definition 1.1. Definition 2.1 introduces Vietoris-Rips, Čech and Delaunay complexes on a finite set A in any metric space M or for $A \subset \mathbb{R}^N$ for Delaunay complexes.

Let M be any metric space with a distance d satisfying all metric axioms. An example of a metric space is $\mathbb{R}^N$ with the Euclidean metric. For any points $p, q \in A \subset M$, the edge $e = [p, q]$ has the length $d(p, q)$. For a point cloud $A \subset \mathbb{R}^N$, $e = [p, q]$ has the Euclidean length $|p - q|$ and can be geometrically interpreted as the straight-line segment connecting the points $p, q \in A \subset \mathbb{R}^N$.

Definition 2.1 introduces the simplicial complexes $\text{VR}(A; \alpha)$ and $\text{Čech}(A; \alpha)$ on any finite set A inside an ambient metric space M, although $A = M$ is possible. For a point $p \in A$ and $\alpha \geq 0$, let $\bar{B}(p; \alpha) \subset M$ denote the closed ball with centre p and radius α. A Delaunay complex $\text{Del}(A; \alpha) \subset \mathbb{R}^N$ will be defined for a finite set A only in $\mathbb{R}^N$ because of extra complications arising if a point set A lives in a more general metric space (Boissonnat et al. 2018).

Definition 2.1 (*Complexes* VR, *Čech, and* Del) Let A be any finite set in a metric space M. Fix a *scale $\alpha \geq 0$*. Each complex $C(A; \alpha)$ below has the vertex set A.

(a) The *Vietoris-Rips* complex $\text{VR}(A; \alpha)$ has all simplices on points $p_1, \ldots, p_k \in A$ whose pairwise distances are at most 2α, so $d(p_i, p_j) \leq 2\alpha$ for $i \neq j$ in $\{1, \ldots, k\}$.
(b) The *Čech* complex $\text{Čech}(A; \alpha)$ has all simplices on points $p_1, \ldots, p_k \in A$ such that the full intersection $\cap_{i=1}^{k} \bar{B}(p_i; \alpha)$ is not empty.
(c) For any finite set of points $A \subset \mathbb{R}^N$, the convex hull of A is the intersection of all closed half-spaces of $\mathbb{R}^N$ containing A. Each point $p_i \in A$ has the *Voronoi domain*

$$V(p_i) = \{q \in \mathbb{R}^N \mid |q - p_i| \leq |q - p_j| \text{ for any point } p_j \in A, \ p_j \neq p_i\}.$$

The *Delaunay* complex Del$(A; \alpha)$ has all simplices on points $p_1, \ldots, p_k \in A$ such that $\cap_{i=1}^{k}(V(p_i) \cap \bar{B}(p_i; \alpha)) \neq \emptyset$ (Delaunay 1934). Alternatively, a simplex σ on points $p_1, \ldots, p_k \in A$ is called a *Delaunay* simplex if there is an $(N-1)$-dimensional sphere S^{N-1} that passes through the points $p_1, \ldots, p_k$ and does not enclose any points of A (Shewchuk 2000).

In a degenerate case, the smallest $(k-2)$-dimensional sphere S^{k-2} above can contain more than k points of A. If σ is enlarged to the convex hull H of all points in $A \cap S^{k-2}$, then Del$(A; \alpha)$ becomes a polyhedral *Delaunay mosaic*.

For simplicity, we can choose any triangulation of H into Delaunay simplices. When the scale α becomes too large, Del$(A; \alpha) \subset \mathbb{R}^N$ stops growing and becomes a *Delaunay triangulation* of the convex hull of A, which is unique in general position. The complexes of the types above will be called *geometric* complexes for brevity. ∎

Both complexes VR$(A; \alpha)$ and Čech$(A; \alpha)$ are abstract and so are not embedded in $\mathbb{R}^N$, even if $A \subset \mathbb{R}^N$. Though Del$(A; \alpha)$ is embedded into $\mathbb{R}^N$, its construction has a near-linear time or quadratic time in the size of A only in dimensions $N = 2, 3$ (Cignoni et al. 1998). For high dimensions $N > 3$ or any metric space M, the simplest complex to build and store is VR$(A; \alpha)$. Indeed, the Vietoris-Rips complex VR$(A; \alpha)$ is a flag complex determined by its 1-dimensional skeleton VR$^1(A; \alpha)$ so that any simplex of VR$(A; \alpha)$ is built on a complete subgraph whose vertices are pairwise connected by edges in VR$^1(A; \alpha)$.

The key idea of persistence is to view any point cloud $A \subset \mathbb{R}^N$ through the lens of a variable scale $\alpha \geq 0$. When the scale α is increasing from the initial value 0, we can form a new topological space from A by replacing points with closed balls of a radius α. Then persistent homology identifies topological features of these spaces that 'persist' over a long interval of the scale α.

More formally, for any fixed scale $\alpha \geq 0$, the union $\cup_{p \in A} \bar{B}(p; \alpha)$ of closed balls is homotopy equivalent to the Čech complex Čech$(A; \alpha)$ and also to the Delaunay complex Del$(A; \alpha) \subset \mathbb{R}^N$ by the Nerve Lemma (Hatcher 2002, Corollary 4 G.3), see also the persistent version in (Chazal and Oudot 2008, Lemma 3.4).

For any geometric complex $C(A; \alpha)$ from Definition 2.1, all connected components of $C(A; \alpha)$ are in a 1–1 correspondence with all connected components of the union $\cup_{p \in A} \bar{B}(p; \alpha)$ of the closed balls centred at all $p \in A$. If an edge $e = [p, q]$ enters a complex $C(A; \alpha)$ at a scale α, then $\alpha = d(p, q)/2$.

Definition 2.2 makes sense for any filtration of simplicial complexes from Definition 1.1, not only for geometric complexes from Definition 2.1.

Definition 2.2 (*Short, medium, long edges in a filtration*) Let $\{C(A; \alpha)\}$ be any filtration of complexes on a finite vertex set A, see Definition 1.1. Let an edge $e = [p, q]$ between points $p, q \in A$ enter the complex $C(A; \alpha)$ at the scale $\alpha = d(p, q)/2$.
(a) Consider the 1-dimensional graph $C'(A; \alpha)$ with vertex set A and all edges from $C(A; \alpha)$ except the edge e. If the endpoints of e are in different connected components of $C'(A; \alpha)$, then the edge e is called *short* in the filtration $\{C(A; \alpha)\}$.
(b) The edge e is called *long* in $\{C(A; \alpha)\}$ if A has a vertex v such that $C(A; \alpha)$ has the 2-simplex $\triangle pqv$ and both edges $[p, v], [v, q]$ are in $C(A; \alpha')$ for some $\alpha' < \alpha$.
(c) If e is neither short nor long, then the edge e is called *medium* in $\{C(A; \alpha)\}$. ∎

Definition 2.2(b) implies that any long edge enters $C(A; \alpha)$ with a 2-simplex $\triangle pqv$ at the same scale α and the boundary of this 2-simplex is homologically trivial in $C(A; \alpha)$ due to the other two edges $[p, v]$ and $[v, q]$ that entered the filtration at a smaller scale $\alpha' < \alpha$.

Lemma 2.3 (Classes of edges) *For any finite set A and a filtration $\{C(A; \alpha)\}$ from Definition 2.1, all edges are split into disjoint classes: short, medium, long.* ∎

Proof By Definition 2.2(b), the endpoints p, q of any long edge $e = [p, q] \subset C(A; \alpha)$ are connected by a chain of two edges $[p, v] \cup [v, q]$ that entered the filtration at a smaller scale $\alpha' < \alpha$. Hence the long edge e cannot be short by Definition 2.2(a). So the three classes of edges in Definition 2.2 are disjoint. □

Definition 2.2 defined classes of edges for any filtration of complexes. Proposition 2.4 interprets long edges in VR and Cech filtrations via distances.

Proposition 2.4 (Long edges in VR, Čech, Del) *Let A be a finite metric space.*
(a) An edge $e = [p, q]$ in the Vietoris-Rips filtration $\{VR(A; \alpha)\}$, is long if and only if A has a point v such that $e = [p, q]$ is strictly longest in the 2-simplex $\triangle pqv$.
(b) An edge $e = [p, q]$ in the Čech filtration $\{\check{C}ech(A; \alpha)\}$ is long if and only if A has a point v such that $e = [p, q]$ is strictly longest in the 2-simplex $\triangle pqv$ and the triple intersection $\bar{B}(p; \alpha) \cap \bar{B}(q; \alpha) \cap \bar{B}(v; \alpha)$ is not empty for $\alpha = d(p, q)/2$.
(c) For $A \subset \mathbb{R}^N$, an edge $e = [p, q]$ in the Delaunay filtration $\{Del(A; \alpha)\}$ is long if and only if A has a point v such that $e = [p, q]$ is strictly longest in the 2-simplex $\triangle pqv$ and $V(p) \cap \bar{B}(p; \alpha) \cap V(q) \cap \bar{B}(q; \alpha) \cap V(v) \cap \bar{B}(v; \alpha) \neq \emptyset$ for $\alpha = |p - q|/2$.
(d) For $A \subset \mathbb{R}^N$ and any filtration from Definition 2.1 and an edge $[p, q]$ in $C(A; \alpha)$, if A has a point v such that the angle at v in $\triangle pqv$ is not acute, then $[p, q]$ is long. ∎

Proof For all filtrations from Definition 2.1, an edge e enters $C(A; \alpha)$ at the scale $\alpha = d(p, q)/2$. By Definition 2.2(b), a long edge enters $C(A; \alpha)$ together with a 2-simplex $\triangle pqv$ for some $v \in A$, while the other two edges $[p, v]$, $[v, q]$ entered the filtration at a smaller scale α', the edge $e = [p, q]$ is longest in the 2-simplex $\triangle pqv$.
(b) For the Čech filtration, the triple intersection $\bar{B}(p; \alpha) \cap \bar{B}(q; \alpha) \cap \bar{B}(v; \alpha)$ is non-empty to guarantee that $\check{C}ech(A; \alpha)$ includes $\triangle pqv$ by Definition 2.1(b).
(c) For the Delaunay filtration, $V(p) \cap \bar{B}(p; \alpha) \cap V(q) \cap \bar{B}(q; \alpha) \cap V(v) \cap \bar{B}(v; \alpha)$ is not empty to guarantee that $Del(A; \alpha)$ includes $\triangle pqv$ by Definition 2.1(c).
(d) For all filtrations (Vietoris-Rips, Čech and Delaunay) and $A \subset \mathbb{R}^N$, if the angle at v in $\triangle pqv$ is not acute, then $[p, q]$ is strictly longest in $\triangle pqv$, which finishes the proof for the Vietoris-Rips filtration by part (a). The closed ball $\bar{B}(u; \alpha)$ centred at the mid-point u of $[p, q]$ contains all p, q, v, so the point u belongs to $\bar{B}(p; \alpha) \cap \bar{B}(q; \alpha) \cap \bar{B}(v; \alpha)$, which finishes the proof for the Čech filtration by part (b).

For the Delaunay filtration, since the edge $[p, q]$ entered $Del(A; \alpha)$ at the scale α, Definition 2.1(c) gives an $(N - 1)$-dimensional sphere S^{N-1} that passes through p, q and does not enclose any point of A. Let $S(v)$ be the smallest $(N - 1)$-dimensional sphere that passes through p, q, v. If $S(v)$ encloses (strictly inside) no points of A, then the 2-simplex $\triangle pqv$ is Delaunay by Definition 2.1(c) and enters $Del(A; \alpha)$ together with $[p, q]$ at $\alpha = |p - q|/2$, so $[p, q]$ is long by Definition 2.2(b). Otherwise, we will find another empty sphere circumscribing a non-acute Delaunay 2-simplex on $[p, q]$.
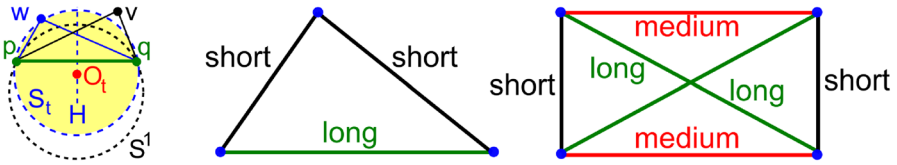
Fig. 3 Left: an edge $[p, q]$ opposite to a non-acute angle in a 2-simplex $\triangle pqw$, see the proof of Proposition 2.4(d). **Middle** and **Right**: classes of edges by Definition 2.2 in Example 2.5

The centres of the spheres S^{N-1} and $S(v)$ lie in the $(N-1)$-dimensional hyperspace H that perpendicularly splits the edge $[p, q]$ at its mid-point u. Connect these centres by the straight-line path of points O_t, $t \in [0, 1]$, within H. For every centre $O_t \in H$, consider the $(N-1)$-dimensional sphere S_t with the radius $R_t = |O_t - p| = |O_t - q|$ so that S_t passes through p, q for all $t \in [0, 1]$, see Fig. 3(left).

Then the continuous family of spheres deforming from S^{N-1} to $S(v)$ should contain a sphere S_t that passes through a point $w \in A - \{p, q\}$ and encloses no points of A. This point w should lie inside the spherical segment bounded by $S(v)$ and the $(N-1)$-dimensional hyperspace H_1 passing through $[p, q]$ orthogonally to $[u, O_1]$.

Since this segment is not larger than a half-ball bounded by $S(v)$, any such point w has a non-acute angle $\angle pwq$ on the diameter $[p, q]$ of the $(N-2)$-dimensional sphere $S(v) \cap H_1$. Then the non-acute 2-simplex $\triangle pqw$ is Delaunay by Definition 2.1(c) and enters $Del(A; \alpha)$ together with $[p, q]$, so $[p, q]$ is long by Definition 2.2(b). □

Example 2.5 (Classes of edges on 3 and 4 points) For any 3-point set $A \subset \mathbb{R}^N$, let the edges of A have lengths $|e_1| \leq |e_2| < |e_3|$. By Definition 2.2, in $\{VR(A; \alpha)\}$ the edge e_3 is long whilst the edges e_1, e_2 are short, see Fig. 3 (middle). If $|e_1| < |e_2| = |e_3|$, then the edge e_1 is short but both edges e_2, e_3 are medium, not long. If $|e_1| = |e_2| = |e_3|$, then all three edges are medium. Let $C(A; \alpha)$ be any geometric complex from Definition 2.1 on a finite set $A \subset \mathbb{R}^2$. If the set A consists of four vertices of the unit square, all square sides are medium whilst both diagonals are long, see Fig. 3 (right). If the set A consists of four vertices of a rectangle that is not a square, the two shorter sides are short, the longer sides are medium and both diagonals are long. ∎

3 Tails without medium edges in a metric space

As usual, we consider homology groups with coefficients in a field, say $\mathbb{Z}_2$.

Proposition 3.1 (No medium edges $\Rightarrow$ trivial H_1) *For any filtration $\{C(A; \alpha)\}$ on a finite set A from Definition 1.1, when a scale $\alpha \geq 0$ is increasing, a new homology cycle in $H_1(C(A; \alpha))$ can be created only due to a medium edge in $C(A; \alpha)$. Hence, if $\{C(A; \alpha)\}$ has no medium edges, then $H_1(C(A; \alpha))$ is trivial for $\alpha \geq 0$.* ∎

Proof When building the complex $C(A; \alpha)$, if we add a short edge e, by Definition 2.2(a), the previously disjoint components of $C^1(A; \alpha)$ containing the endpoints p, q of e become connected. Hence no 1-dimensional cycle in $C^1(A; \alpha)$ is created.

For any α, let a cycle γ have just appeared in $H_1(C(A; \alpha))$, represented by several edges including $e_1, ...e_k$ that have appeared at the same scale α. By Lemma 2.3 each
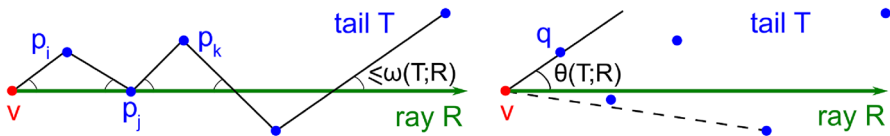
Fig. 4 A tail T around a ray R with vertex v in $\mathbb{R}^2$, see Definitions 3.4 and 3.6. **Left**: all angles are not greater than the angular deviation $\omega(T; R)$. **Right**: the angular thickness $\theta(T; R)$ can be smaller than the angular deviation $\omega(T; R)$

e_i is either short, medium or long. By Definition 2.2(b) any long edge $e = [p, q]$ enters $C(A; \alpha)$ strictly after two shorter edges $[p, v]$, $[v, q]$, and at the same time as the triangle $\triangle pqv$. The cycle γ including the edge $[p, q]$ is homologically equivalent to the cycle with $[p, q]$ replaced with the chain $[p, v] \cup [v, q]$. Hence we can assume that all $e_1, \dots, e_k$ are either short or medium. Since the endpoints of e_i are connected by the complementary path $\gamma - e_i$, each e_i cannot be short by Definition 2.2(a) for $i = 1, \dots, k$. So γ contains at least one medium edge. Since only medium edges lead to non-trivial cycles, if A has no medium edges, then $H_1(C(A; \alpha))$ is trivial.

Definition 3.2 (*Tail of points*) For a fixed filtration $\{C(A; \alpha)\}$ on a finite set A from Definition 1.1, a *tail* T in a metric space M is any ordered sequence $T = \{p_1, \dots, p_n\}$, where p_1 is the *vertex* of T, any edge $[p_i, p_{i+1}]$ between successive points is short, and any edge $[p_i, p_j]$ between non-successive points is long for any $1 \leq i < j \leq n$. ∎

Proposition 3.3 (Tails have trivial PD$_1$) *Any tail T from Definition 3.2 for a filtration $\{C(T; \alpha)\}$ of complexes from Definition 1.1 has trivial 1D persistence.*

Proof Since any tail T has no medium edges by Lemma 2.3, the tail T has trivial $H_1(C(T; \alpha))$ for any $\alpha \geq 0$ by Proposition 3.1, hence trivial 1D persistence. □

If vectors are not explicitly specified, all edges and straight lines are unoriented. We measure the angle between unoriented straight lines as their minimum angle within $[0, \frac{\pi}{2}]$, see Fig. 4(left).

Definition 3.4 (*Angular deviation $\omega(T; R)$ from a ray R*) In $\mathbb{R}^N$, a *ray* is any half-infinite line R going from a point v (the *vertex* of R). For any sequence $T = \{p_1, \dots, p_n\}$ of ordered points in $\mathbb{R}^N$, the *angular deviation* $\omega(T; R)$ of T relative to R is the maximum angle $\angle(R, [p, q]) \in [0, \frac{\pi}{2}]$ over all distinct points $p, q \in T$. ∎

Lemma 3.5 (Tails in $\mathbb{R}^N$) *In $\mathbb{R}^N$, let R be a straight infinite ray with a vertex $v = p_1$ and T be any sequence of points $p_1, \dots, p_n$ with an angular deviation $\omega(T; R) < \frac{\pi}{4}$. (a) For any $i < j < k$, the angle $\angle p_i p_j p_k$ is non-acute. The edge between the non-successive points p_i, p_k is long in any filtration $\{C(T; \alpha)\}$ in Definition 2.1. (b) Any edge between successive points p_{j-1}, p_j, $j = 2, \dots, n$, is short in $\{C(T; \alpha)\}$. Hence T has no medium edges in $\{C(T; \alpha)\}$ and is a tail by Definition 3.2.*

Proof (a) The condition $\omega(T; R) < \frac{\pi}{4}$ implies that all points of T can be ordered by their distance from the vertex $v = p_1$ to their orthogonal projections in the ray R.

Apply a parallel shift to p_i, p_j, p_k so that $p_j \in R$. In the 2-simplex $\triangle p_i p_j p_k$, the angle

$$\angle p_i p_j p_k = \pi - \angle(R, [p_j p_i]) - \angle(R, [p_j p_k]) \geq \pi - 2\omega(T; R) > \frac{\pi}{2}$$

is non-acute due to $\omega(T; R) < \frac{\pi}{4}$, hence strictly largest. By Proposition 2.4(d) the edge $[p_i, p_k]$ is long in any filtration $\{C(T; \alpha)\}$ in the sense of Definition 2.2(b). In particular, the edge $[p_i, p_k]$ is longer than both $[p_i, p_j]$ and $[p_j, p_k]$ for any $i < j < k$. **(b)** The points p_{j-1}, p_j remain in disjoint components of $C^1(T; \alpha)$ after adding all other edges of the same length $|p_j - p_{j-1}|$. Indeed, we proved above that any other edge connecting non-successive points p_i, p_k for $i \leq j - 1 < j \leq k$ is longer than the edge $[p_{j-1}, p_j]$ between intermediate successive points. $\qquad\square$

Figure 4 (right) illustrates the angular thickness below for Theorem 4.4 later.

Definition 3.6 (*Angular thickness θ*) Let $R \subset \mathbb{R}^N$ be a ray with a vertex $v = p_1$, $T = \{p_1, \ldots, p_n\}$ be a finite sequence of points. The *angular thickness* $\theta(T; R)$ of T with respect to R is the maximum angle $\angle(R, [p_1, p_i])$ for $i = 2, \ldots, n$. $\qquad\blacksquare$

4 Persistence for long wedges and with tails

This section proves main Theorem 4.4 saying that the 1D persistence for a point cloud A remains unchanged under adding a suitable tail T of points to A. The key step is Theorem 4.2 describing how to compute the 1D persistence for a union of point clouds $\cup_{i=1}^k A_i$ sharing a single point as defined below.

Definition 4.1 (*A long wedge*) Let $A_1, \ldots, A_k$ be finite point clouds sharing one common point v. In a filtration $\{C(\cup_{i=1}^k A_i; \alpha)\}$ from Definition 1.1, call a simplex *heterogeneous* if its vertices don't include v and belong to at least two different clouds A_i for $i = 1, \ldots, k$. Assume that any heterogeneous edge of $\{C(\cup_{i=1}^k A_i; \alpha)\}$ is long in the sense of Definition 2.2(b). Also assume that if any heterogeneous 2-simplex abc enters the filtration $\{C(\cup_{i=1}^k A_i; \alpha)\}$ at a scale α, then $C(\cup_{i=1}^k A_i; \alpha)$ includes the 2-simplices abv, bcv, cav. Then the union $\cup_{i=1}^k A_i$ is called a *long wedge*. $\qquad\blacksquare$

In topology, a *wedge* (or bouqet) $\vee_{i=1}^k C(A_i; \alpha)$ of complexes, each with a base point v_j, is the quotient of the disjoint $\sqcup_{i=1}^k C(A_i; \alpha)$, where all base points $v_1, \ldots, v_k$ are collapsed to one point v. (Hatcher 2002, Corollary 2.25) proves an isomorphism $H_1(\vee_{i=1}^k C(A_i; \alpha)) \rightarrow \oplus_{i=1}^k H_1(C(A_i; \alpha))$. Theorem 4.2 proves a similar isomorphism for the larger complex $C(\cup_{i=1}^k A_i; \alpha)$ of a long wedge $\cup_{i=1}^k A_i$ of point clouds instead of the wedge $\vee_{i=1}^k C(A_i; \alpha)$ of smaller complexes.

Theorem 4.2 (Persistence of a long wedge) *For any filtration $\{C(\cup_{i=1}^k A_i; \alpha)\}$ of a long wedge from Definition 4.1, $H_1(C(\cup_{i=1}^k A_i; \alpha))$ is isomorphic to the direct sum $\oplus_{i=1}^k H_1(C(A_i; \alpha))$ for all α. Hence the 1D persistence diagram $\mathrm{PD}_1\{C(\cup_{i=1}^k A_i; \alpha)\}$ is the union of the 1D persistence diagrams $\mathrm{PD}_1\{C(A_i; \alpha)\}$ for $i = 1, \ldots, k$.* $\qquad\blacksquare$

Proof Due to the isomorphism $H_1(\vee_{i=1}^k C(A_i; \alpha)) \cong \oplus_{i=1}^k H_1(C(A_i; \alpha))$ by (Hatcher 2002, Corollary 2.25), it suffices to prove that $H_1(\vee_{i=1}^k C(A_i; \alpha)) \cong H_1(C(\cup_{i=1}^k A_i; \alpha))$.

The inclusion $\vee_{i=1}^k C(A_i; \alpha) \subset C(\cup_{i=1}^k A_i; \alpha)$ induces the homomorphism $H_1(\vee_{i=1}^k C(A_i; \alpha)) \to H_1(C(\cup_{i=1}^k A_i; \alpha))$ whose bijectivity is proved below.

Surjectivity of h. By Definition 2.2(b) any long edge $e = [p, q]$ belongs to a complex $C(\cup_{i=1}^k A_i; \alpha)$ together with a 2-simplex pvq whose edges $[p, v]$ and $[q, v]$ have already entered $C(\cup_{i=1}^k A_i; \alpha')$ for some $\alpha' < \alpha$.

Replace the edge $[p, q]$ with the homologous chain $[p, v] \cup [v, q]$ in $C(\cup_{i=1}^k A_i; \alpha)$. Continue applying these replacements for other long edges until any cycle of edges in $C(\cup_{i=1}^k A_i; \alpha)$ becomes homologous to a sum of non-long edges.

By Definition 4.1, both endpoints of any remaining non-long edge in $C(\cup_{i=1}^k A_i; \alpha)$ belong to the same cloud A_i. Then the resulting cycle is a sum of k sums $s_1, \ldots, s_k$, where each s_i is a sum of only edges from $C(A_i; \alpha)$. Since all clouds A_i share a single point, the resulting cycle is a wedge (1-point union) of the sums $s_1, \ldots, s_k$, which should be cycles in $C(A_i; \alpha)$ for $i = 1, \ldots, k$, respectively. So any cycle in $H_1(C(\cup_{i=1}^k A_i; \alpha))$ is homologous to an element in $H_1(\vee_{i=1}^k C(A_i; \alpha))$.

Injectivity of h. It remains to prove that if any 1-dimensional cycle γ in $\vee_{i=1}^k C(A_i; \alpha)$ is bounded by a 2-dimensional chain $\sigma \in C(\cup_{i=1}^k A_i; \alpha)$, then γ is bounded by a chain τ in $\vee_{i=1}^k C(A_i; \alpha)$. By Definition 4.1 replace any heterogeneous 2-simplex $[abc]$ in the closure of $C(\cup_{i=1}^k A_i; \alpha) - (\vee_{i=1}^k C(A_i; \alpha))$ with the sum of non-heterogeneous simplices $[abv] + [bcv] + [cav]$, whose total boundary is $\partial[abc]$.

After all such replacements, we get a chain τ that has the same boundary $\partial \tau = \gamma$ and has no heterogeneous simplices. The boundary $\partial \tau$ also has no heterogeneous edges $[p, q]$ with $p \in A_i - \{v\}$ and $q \in A_j - \{v\}$ for $i \neq j$, else $\gamma = \partial \tau$ is not within the wedge $\vee_{i=1}^k C(A_i; \alpha)$ of complexes. Hence every 2-simplex of τ is within a single cloud A_i for some $i = 1, \ldots, k$, so the whole chain τ is within $\vee_{i=1}^k C(A_i; \alpha)$. □

Definition 4.3 is needed by (Bauer and Edelsbrunner 2017, Theorem 5.10) to guarantee that the filtration of Čech and Delaunay complexes have the same persistence.

Definition 4.3 (*A cloud in general position*) A finite cloud $A \subset \mathbb{R}^N$ is in *general position* if every subset $P \subset A$ of at most $N + 1$ points is affinely independent, and no point of $A - P$ lies on the smallest $(N-1)$-dimensional circumsphere of P. ∎

Theorem 4.4 can be considered a Euclidean example of Theorem 4.2 and describes sufficient conditions for a cloud A and a tail T to guarantee that three types of filtrations on $A \cup T$ and A have the same persistence PD_1.

Theorem 4.4 (A long wedge with a tail) *Let $A \subset \mathbb{R}^N$ be a finite set, $v \in A$ be on the boundary of the convex hull of A, and R be a ray with a vertex v so that $\mu(R; A) = \min_{p \in A - \{v\}} \angle(R, [v, p]) \geq \frac{\pi}{2}$. Let T be a tail with the vertex v such that $\mu(R; A) \geq \theta(T; R) + \frac{\pi}{2}$ and $A \cup T$ is in general position by Definition 4.3. For any filtration from Definition 2.1, we have that $\text{PD}_1\{C(A \cup T; \alpha)\} = \text{PD}_1\{C(A; \alpha)\}$.* ∎
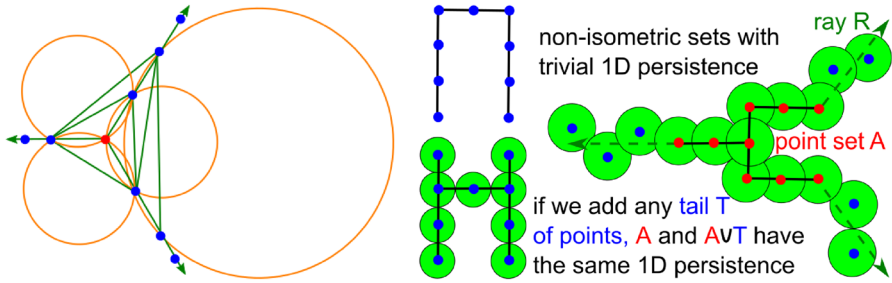
Fig. 5 Left: the cloud A in Theorem 4.4 can be a single red point extendable by tails of blue points along straight rays that form non-acute angles. Then all Delaunay triangles are obtuse, circumscribed by orange circles, meaning that $PD_1\{Del(C; \alpha)\} = \emptyset$. **Right**: a tail T can be generically perturbed under conditions of Theorem 4.4 without changing PD_1 (colour figure online)

Proof Any heterogeneous edge $[p, q]$ with $p \in A$ and $q \in T$ has a non-acute angle at v

$$\angle pvq \geq \angle(R, [v, p]) - \angle(R, [v, q]) \geq \mu - \angle(R, [v, q]) \geq \mu - \theta(T; R) > \frac{\pi}{2}.$$

Due to the point v, the heterogeneous edge $[p, q]$ is long in $\{C(A \cup T; \alpha)\}$ by Proposition 2.4(d). To prove that $A \cup T$ is a long wedge by Definition 4.1, consider any heterogeneous 2-simplex abc in the complex $\{C(A \cup T; \alpha)\}$. In the boundary $\partial[abc]$, any heterogeneous edge, say $[a, b]$, is strictly the longest by the argument above, while the edges $[a, v]$, $[b, v]$ are no longer heterogeneous, so $A \cup T$ is a long wedge.

In the case of a Čech filtration, let $\triangle abc$ be any heterogeneous 2-simplex in $\check{C}ech(A \cup T; \alpha)$ such that (say) $a \in A$ and $b, c \in T$. For the heterogenous edges $[a, b]$ and $[a, c]$, the earlier proved inequalities $\angle avb \geq \frac{\pi}{2}$ and $\angle avc \geq \frac{\pi}{2}$ implies that v belongs to the smallest closed circumballs of $[a, b]$ and $[a, c]$, hence to the smallest closed circumball of $\triangle abc$. Then the 3-simplex $abcv$ and all its faces belong to $\check{C}ech(A \cup T; \alpha)$. All conditions of Definition 4.1 hold, so $A \cup T$ is a long wedge.

Since the tail T has the trivial (empty) 1D persistence by Proposition 3.3, Theorem 4.2 implies that $PD_1\{C(A \cup T; \alpha)\} = PD_1\{C(A; \alpha)\}$ for any filtration form Definition 2.1. By (Bauer and Edelsbrunner 2017, Theorem 5.10), the Delaunay and Čech filtrations have the same persistence for clouds in general position, which finishes the proof. □

Figure 5 (left) illustrates a Delaunay filtration on a cloud $A \subset \mathbb{R}^2$. All blue points lie on rays that have pairwise angles $120°$ and emanate from a red point v so that all green Delaunay triangles are obtuse with all orange circumcircles not enclosing any points of A, which implies that $PD_1\{C(A; \alpha)\}$ is empty.

Corollary 4.5 *(Clouds with $PD_1 = \emptyset$) If a point cloud A has $PD_1\{C(A; \alpha)\} = \emptyset$, then any long wedge $A \cup T$ with a tail T has $PD_1\{C(A \cup T; \alpha)\} = \emptyset$.* ∎

Proof Since the tail T has trivial 1D persistence by Proposition 3.3, Theorem 4.2 implies that $PD_1\{C(A \cup T; \alpha)\} = PD_1\{C(A; \alpha)\} = \emptyset$, see Fig. 5 (right). □
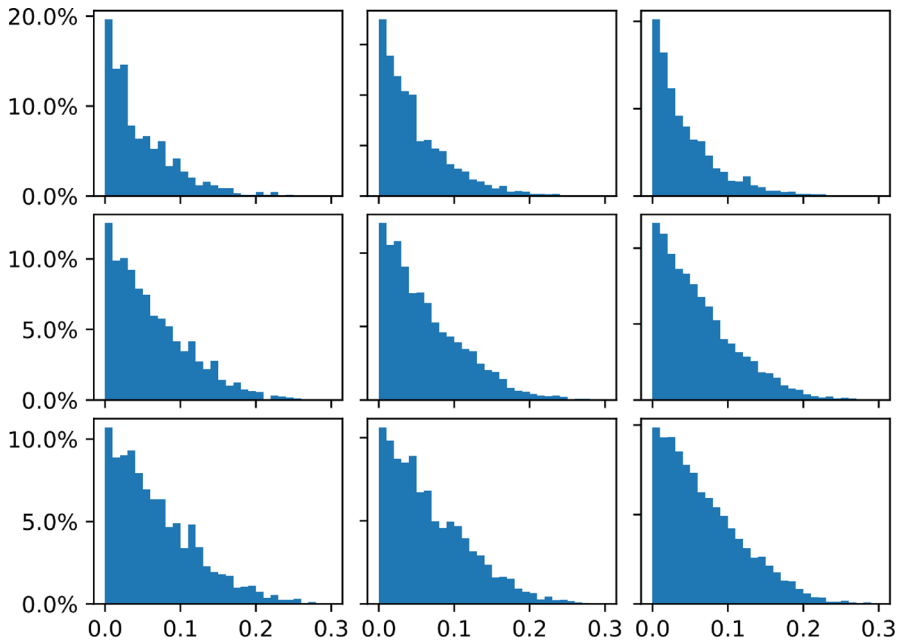
Fig. 6 Histograms of the persistence $p = \text{death} - \text{birth}$ in 1000 point sets in nine configurations of the parameters n and N. The x-axis is the persistence p, the y-axis is the percentage of pairs (birth,death) with the given persistence p. Top row: $N = 2$; middle row: $N = 5$; bottom row: $N = 8$. Left column: $n = 10$; middle column: $n = 15$; right column $n = 20$

5 Experiments on persistence of random sets

The experiments in this section use the Vietoris-Rips filtration whose 1-dimensional persistence is computed by Bauer (2021), a fast implementation of Vietoris-Rips persistence. The code of the first author is available in Smith (2022).

The aim is to understand how often random point sets have trivial persistence or cycles with only low persistence, see more general conjectures (Bobrowski and Skraba 2022). The experiments depend on two parameters, the size n of a set, and the dimension N that the point set lies in. For each n, N in the ranges chosen, we generate 1000 point sets of n points uniformly sampled in a unit N-dimensional cube.

Figure 6 shows histograms of the 1-dimensional persistence (death−birth) for nine configurations of the parameters: set sizes $n = 10, 15, 20$ and dimensions $N = 2, 5, 8$. Each histogram highlights that one-dimensional persistent features are skewed towards a low persistence. Geometrically, the pairs (birth,death) would be close to the diagonal in a persistence diagram.

Recall that highly persistent features (birth,death) are naturally separated from others with lower persistence $p = \text{death} - \text{birth}$ by the widest diagonal gap in the persistence diagram, see (Smith and Kurlin 2021). If we order all pairs (birth,death) by their persistence $0 < p_1 \leq \cdots \leq p_k$, the widest gap has the largest difference
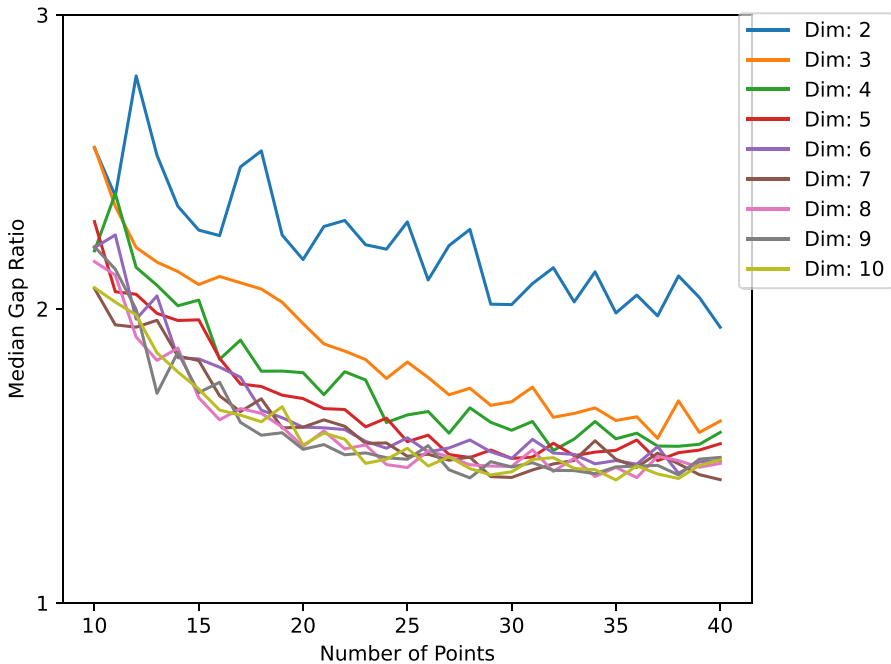
Fig. 7 The median gap ratio of a point set with at least two 1D persistent features, as the set size varies from $n = 10$ to $n = 40$ and the dimension N varies from $N = 2$ to $N = 10$

$p_{i+1} - p_i$ over $i = 1, \ldots, k-1$. This widest gap can separate several pairs (birth,death) from the rest, not necessarily just a single feature.

However, the first widest gap is significant only if it can be easily distinguished from the second widest gap. So the significance of persistence can be measured as the ratio of the first widest gap over the second widest gap. This invariant up to uniform scaling of given data is called the *gap ratio*. Figure 7 shows the median gap ratio calculated over 1000 random point clouds in a unit cube for dimensions $N = 2, \ldots, 10$ and point set sizes $n = 10, \ldots, 40$.

Figure 7 implies that for higher dimensions N, the median gap ratio quickly decreases to within the range [1,2] as the number n of points is increasing. Hence, for pure random clouds when a persistence diagram contains at least two pairs (birth,death) above the diagonal, it is harder to separate highly persistent features from noisy artefacts that are close to the diagonal.

Figure 7 also seems to suggest a limiting distribution as $N \to +\infty$.

6 Conclusions and discussion of other invariants

Main Theorem 4.4 showed how one can add an arbitrarily long tails to an existing point set without affecting the 1-dimensional persistent homology. Corollary 4.5 implies that

families of sets with trivial 1D persistence form vast continuous subspaces in the space of isometry classes of finite sets.

The bottleneck distance between persistence diagrams vanishes on these subspaces and cannot have a lower bound. We conjecture that Theorem 4.4 extends to any higher-dimensional persistence in the following open problem.

Problem 6.1 (*Adding tails preserves any persistence*) Check if, for any point cloud $A \subset \mathbb{R}^N$ and a tail satisfying Theorem 4.4, adding the tail T to A preserves any k-dimensional persistence, so $\mathrm{PD}_k\{C(A \cup T; \alpha)\} = \mathrm{PD}_k\{C(A; \alpha)\}$ for $k \geq 1$. ∎

Theorem 4.4 gave only sufficient conditions that guarantee the same 1D persistence under adding a tail. Problem 6.2 asks to weaken these conditions.

Problem 6.2 (*Necessary conditions for preserving persistence*) For each filtration from Definition 2.1, find sufficient a necessary conditions on a cloud A and its tail T such that $\mathrm{PD}_k\{C(A \cup T; \alpha)\} = \mathrm{PD}_k\{C(A; \alpha)\}$ for $k \geq 1$. ∎

Oudot and Solomon (2020) previously asked to find one point cloud for a given persistence: "If a given persistence module does come from a point cloud, can that point cloud be computed effectively?" Corollary 4.5 described a generic family of clouds $A \cup T$ that all have trivial persistence $\mathrm{PD}_1 = \emptyset$. The deeper problem below requires us to geometrically interpret persistence as an equivalence of clouds.

Problem 6.3 (*Persistence as equivalence*) Geometrically describe an equivalence relation on point clouds A. e.g. as transformations of the ambient space, whose classes are in a 1–1 correspondence with persistence diagrams $\mathrm{PD}_k\{C(A; \alpha)\}$ for $k \geq 1$. ∎

Theorem 4.4 motivated comparisons of persistent homology with other isometry invariants of point clouds. For finite sets of m ordered points, a complete isometry invariant is a classical distance $m \times m$ matrix whose brute-force adaptation to unlabelled points requires $m!$ permutations. The simpler collection of $\dfrac{m(m-1)}{2}$ pairwise distances (with repetitions) between m unlabeled points is complete for sets in general position (Boutin and Kemper 2004) but do not distinguish infinitely many non-isometric m-point sets for $m \geq 4$.

The local distribution of distances (Mémoli 2011) was recently studied under the name of the Pointwise Distance Distribution (PDD) for finite and periodic sets (Widdowson and Kurlin 2022). The completeness of the PDD is easy for finite sets in general position in $\mathbb{R}^N$ (Widdowson et al. 2022, Theorem 16) and was recently extended to the much harder periodic case (Widdowson and Kurlin 2022, Theorem 4.4). The PDD is conjectured to be complete for $N = 2$ but cannot distinguish counter-examples (Pozdnyakov et al. 2020) for $N = 3$, which were classified by higher order invariants in appendix C of the first version of Widdowson and Kurlin (2021) in 2021.

The recent even stronger invariants (Kurlin 2022, 2024) were proved to be Lipschitz continuous (Kurlin 2023) and complete under rigid motion in any Euclidean space $\mathbb{R}^N$ (Widdowson and Kurlin 2023), extended to metric spaces with measures (Kurlin 2023). The Lipschitz continuity is important for accurate predictions of material properties (Ropers et al. 2022; Balasingham et al. 2024a, b).

Another advantage of the PDD is its near-linear time based on a new algorithm for nearest neighbours (Elkin and Kurlin 2023), which corrected gaps in the past proofs for cover trees (Elkin and Kurlin 2022). The actual speed is so fast that more than 200 billion pairwise comparisons of all 660K+ periodic crystals in the world's largest database of real materials were done within two days on a modest desktop. This experiment detected physically impossible isometric duplicates whose underlying publications are investigated by five journals for data integrity (Widdowson et al. 2022, section 7).

More importantly, the above experiment justified the Crystal Isometry Principle (CRISP) saying that all real periodic crystals have unique locations determined by their complete isometry invariants in a common Crystal Isometry Space continuously parametrised by complete isometry invariants. Even if examples of periodic sets with the same PDD emerge, the slower isoset invariant is provably complete (Anosova and Kurlin 2021) and has continuous metrics (Anosova and Kurlin 2022).

Author Contributions The second author proved the theoretical results and wrote the paper. The first author completed the experiments in section 5 and reviewed the writing.

Declarations

Conflict of interest The authors state that there is no Conflict of interest.

References

Anosova, O., Kurlin, V.: Algorithms for continuous metrics on periodic crystals. arxiv:2205.15298 (2022)
Anosova, O., Kurlin, V.: An isometry classification of periodic point sets. In: Lecture Notes in Computer Science (Proceedings of DGMM), vol. 12708, 229–241 (2021)
Balasingham, J., Zamaraev, V., Kurlin, V.: Accelerating material property prediction using generically complete isometry invariants. Sci. Rep. **14**, 10132 (2024)
Balasingham, J., Zamaraev, V., Kurlin, V.: Material property prediction using graphs based on generically complete isometry invariants. Integr. Mater. Manuf. Innov. (2024). https://doi.org/10.1007/s40192-024-00351-9
Barannikov, S.: The framed Morse complex and its invariants. Adv. Soviet Math. **21**, 93–116 (1994)

Bauer, U.: Ripser: efficient computation of vietoris-rips persistence barcodes. J. Appl. Comput. Geom. **5**(3), 391–423 (2021)

Bauer, U., Edelsbrunner, H.: The Morse theory of Cech and Delaunay complexes. Trans. Am. Math. Soc. **369**(5), 3741–3762 (2017)

Bobrowski, O., Skraba, P.: On the universality of random persistence diagrams. arXiv:2207.03926 (2022)

Boissonnat, J.-D., Dyer, R., Ghosh, A., Martynchuk, N.: An obstruction to Delaunay triangulations in Riemannian manifolds. Discrete Comput. Geom. **59**(1), 226–237 (2018)

Boutin, M., Kemper, G.: On reconstructing n-point configurations from the distribution of distances or areas. Adv. Appl. Math. **32**(4), 709–735 (2004)

Carlsson, G.: Topology and data. Bull. Am. Math. Soc. **46**(2), 255–308 (2009)

Catanzaro, M.J., Curry, J.M., Fasy, B.T., Lazovskis, J., Malen, G., Riess, H., Wang, B., Zabka, M.: Moduli spaces of Morse functions for persistence. J. Appl. Comput. Topol. **4**(3), 353–385 (2020)

Chazal, F., Oudot, S.Y.: Towards persistence-based reconstruction in Euclidean spaces. In: Symposium on Computational Geometry, 232–241 (2008)

Chazal, F., De Silva, V., Glisse, M., Oudot, S.: The Structure and Stability of Persistence Modules. Springer, Berlin (2016)

Cignoni, P., Montani, C., Scopigno, R.: Dewall: A fast divide and conquer delaunay triangulation algorithm in ed. Comput. Aided Des. **30**(5), 333–341 (1998)

Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Stability of persistence diagrams. Discrete Comput. Geom. **37**, 263–271 (2005)

Curry, J.: The fiber of the persistence map for functions on the interval. J. Appl. Comput. Topol. **2**(3), 301–321 (2018)

Delaunay, B.: Sur la sphere vide. Izv. Akad. Nauk USSR **7**, 793–800 (1934)

Edelsbrunner, H., Harer, J.: Persistent homology - a survey. Discrete Comput. Geom. **453** (2008)

Edelsbrunner, H., Heiss, T., Kurlin, V., Smith, P., Wintraecken, M.: The density fingerprint of a periodic point set. In: Proceedings of SoCG, 32–13216 (2021)

Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. In: Proceedings 41st Annual Symposium on Foundations of Computer Science, 454–463 (2000)

Elkin, Y., Kurlin, V.: A new near-linear time algorithm for k-nearest neighbor search using a compressed cover tree. In: International Conference on Machine Learning (ICML), 9267–9311 (2023)

Elkin, Y., Kurlin, V.: Counterexamples expose gaps in the proof of time complexity for cover trees introduced in 2006. In: Topological Data Analysis and Visualization (TopoInVis), 9–17 (2022)

Frosini, P., Landi, C.: Size theory as a topological tool for computer vision. Pattern Recognit. Image Anal. **9**(4), 596–603 (1999)

Ghrist, R.: Barcodes: the persistent topology of data. Bull. Am. Math. Soc. **45**(1), 61–75 (2008)

Hatcher, A.: Algebraic Topology. Cambridge University Press, Cambridge (2002)

Kurlin, V.: A computable and continuous metric on isometry classes of high-dimensional periodic sequences. arxiv:2205.04388 (2022)

Kurlin, V.: Simplexwise distance distributions for finite spaces with metrics and measures. arXiv:2303.14161 (2023)

Kurlin, V.: The strength of a simplex is the key to a continuous isometry classification of euclidean clouds of unlabelled points. arXiv:2303.13486 (2023)

Kurlin, V.: Polynomial-time algorithms for continuous metrics on atomic clouds of unordered points. MATCH Commun. Math. Comput. Chem. **91**, 79–108 (2024)

Mémoli, F.: Gromov–Wasserstein distances and the metric approach to object matching. Found. Comput. Math. **11**, 417–487 (2011)

Oudot, S., Solomon, E.: Inverse problems in topological persistence. In: Topological Data Analysis: The Abel Symposium 2018, 405–433 (2020). Springer

Pozdnyakov, S., Willatt, M., Bartók, A., Ortner, C., Csányi, G., Ceriotti, M.: Incompleteness of atomic structure representations. Phys. Rev. Lett. **125**, 166001 (2020)

Robins, V.: Towards computing homology from finite approximations. In: Topology Proceedings, vol. 24, pp. 503–532 (1999)

Ropers, J., Mosca, M.M., Anosova, O.D., Kurlin, V.A., Cooper, A.I.: Fast predictions of lattice energies by continuous isometry invariants of crystal structures. In: International Conference on Data Analytics and Management in Data Intensive Domains, pp. 178–192 (2022)

Shewchuk, J.R.: Sweep algorithms for constructing higher-dimensional constrained Delaunay triangulations. In: Proceedings of SoCG, 350–359 (2000)

Smith, P.: Trivial 1D persistence. https://github.com/Phil-Smith1/Trivial_1D_Persistence (2022)

Smith, P., Kurlin, V.: Skeletonisation algorithms with theoretical guarantees for unorganised point clouds with high levels of noise. Pattern Recogn. **115**, 107902 (2021)

Weinberger, S.: What is... persistent homology? Notices of the AMS **58**(1), 36–39 (2011)

Widdowson, D., Kurlin, V.: Pointwise distance distributions of periodic point sets. arXiv:2108.04798 (2021)

Widdowson, D., Kurlin, V.: Recognizing rigid patterns of unlabeled point clouds by complete and continuous isometry invariants with no false negatives and no false positives. In: Proceedings of CVPR, 1275–1284 (2023)

Widdowson, D., Kurlin, V.: Resolving the data ambiguity for periodic crystals. In: Proceedings of NeurIPS arXiv:2108.04798 (2022)

Widdowson, D., Mosca, M., Pulido, A., Kurlin, V., Cooper, A.: Average minimum distances of periodic point sets. MATCH Commun. Math. Comput. Chem. **87**, 529–559 (2022)