

# Vietoris-Rips Complexes also Provide Topologically Correct Reconstructions of Sampled Shapes<sup>☆</sup>

Dominique Attali<sup>a,\*</sup>, André Lieutier<sup>b,\*</sup>, David Salinas<sup>a,\*</sup>

<sup>a</sup>*Gipsa-lab, Grenoble, France*

<sup>b</sup>*Dassault Système, Aix-en-Provence, France*

---

## Abstract

Given a point set that samples a shape, we formulate conditions under which the Rips complex of the point set at some scale reflects the homotopy type of the shape. For this, we associate with each compact set  $X$  of  $\mathbb{R}^n$  two real-valued functions  $c_X$  and  $h_X$  defined on  $\mathbb{R}_+$  which provide two measures of how much the set  $X$  fails to be convex at a given scale. First, we show that, when  $P$  is a finite point set, an upper bound on  $c_P(t)$  entails that the Rips complex of  $P$  at scale  $r$  collapses to the Čech complex of  $P$  at scale  $r$  for some suitable values of the parameters  $t$  and  $r$ . Second, we prove that, when  $P$  samples a compact set  $X$ , an upper bound on  $h_X$  over some interval guarantees a topologically correct reconstruction of the shape  $X$  either with a Čech complex of  $P$  or with a Rips complex of  $P$ . Regarding the reconstruction with Čech complexes, our work compares well with previous approaches when  $X$  is a smooth set and surprisingly enough, even improves constants when  $X$  has a positive  $\mu$ -reach. Most importantly, our work shows that Rips complexes can also be used to provide shape reconstructions having the correct homotopy type. This may be of some computational interest in high dimensions.

*Keywords:* Shape reconstruction, Rips complexes, clique complexes, Čech complexes, homotopy equivalence, collapses, high dimensions.

---

<sup>☆</sup>This work is partially supported by ANR Project *GIGA* ANR-09-BLAN-0331-01.

\*Corresponding author

*Email addresses:* [dominique.attali@gipsa-lab.grenoble-inp.fr](mailto:dominique.attali@gipsa-lab.grenoble-inp.fr) (Dominique Attali), [andre.lieutier@3ds.com](mailto:andre.lieutier@3ds.com) (André Lieutier), [david.salinas@gipsa-lab.grenoble-inp.fr](mailto:david.salinas@gipsa-lab.grenoble-inp.fr) (David Salinas)

## 1. Introduction

In this paper, we formulate conditions under which the Rips complex of a point set reflects the homotopy type of the shape that the points sample using measures of how far this shape is from being convex.

*Motivation.* The problem of reconstructing shapes from point clouds arises in many fields, including computer graphics and machine learning [3, 23]. Maybe one of the simplest reconstruction method is to output an  $\alpha$ -offset of the sample points, that is, the union of balls centered at the sample with radius  $\alpha$ . Assuming the shape is a smooth manifold [34, 19] or more generally has a positive  $\mu$ -reach [17], it has been proved that this method provides indeed an approximation with the correct homotopy type for a sufficiently dense sample and a suitable value of the offset parameter  $\alpha$ . Topologically, this is equivalent to computing the  $\alpha$ -shape [24, 26] of the sample points, which can be obtained by first building the Delaunay triangulation and then keeping simplices that fit in an empty ball of radius  $\alpha$  or less.

This approach works well for point clouds in three-dimensional space which have Delaunay triangulations of affordable size [5, 6]. But, as the dimension of the ambient space increases, the size of the Delaunay triangulation explodes [1] and other strategies must be found. If the data points lie on a low-dimensional submanifold, it seems reasonable to ask that the result of the reconstruction depends only upon the intrinsic dimension of the data. This motivated de Silva [21] to introduce *witness complexes* and Boissonnat and Ghosh [13] to define *tangential Delaunay complexes*. For medium dimensions, Boissonnat and al. [12] have modified the data structure representing the Delaunay complex and are able to manage complexes of reasonable size up to dimension six in practice. In particular, they avoid the explicit representation of all Delaunay simplices by storing only edges in what they call the *Delaunay graph*, an idea close to that of using Vietoris-Rips complexes developed in this paper.

*Vietoris-Rips complexes.* Given a point set  $P$  and a scale parameter  $\alpha$ , the *Vietoris-Rips complex* is the simplicial complex whose simplices are subsets of points in  $P$  with diameter at most  $2\alpha$ . Rips complexes are examples of *flag complexes*, and as such enjoy the property that a subset of  $P$  belongs to the complex if and only if all its edges belong to the complex. In other words, Rips complexes are completely determined by the graph of their edges. This

compressed form of storage makes Rips complexes very appealing for computations, at least in high dimensions. Recent results study their simplification through homotopy-preserving edge collapses [36, 37] and edge contractions [8]. However, the strategy of using Rips complexes makes sense only if they are able to reflect the topology of the shape that their vertices sample. A closely related family of simplicial complexes are Čech complexes. Specifically, the *Čech complex* of  $P$  at scale  $\alpha$  consists of all simplices spanned by points in  $P$  that fit in a ball of radius  $\alpha$ . The Čech complex of  $P$  at scale  $\alpha$  is homotopy equivalent to the  $\alpha$ -offset of  $P$  and therefore also possesses the ability to reproduce the topology of the shape sampled by  $P$ . This property was used by Chazal and Oudot [20] to extract topological information on the shape from the Rips complex filtration, by interleaving it with the Čech complex filtration and using persistence topology.

The main contribution of this paper is to unveil a more direct relationship between the respective topologies of the Rips complex and the sampled shape. Specifically, we give conditions under which Rips complexes capture the topology of the shape. In a different setting, it has been proved in [30, 31] that the Rips complex of a point set close enough to a Riemannian manifold for the Gromov-Hausdorff distance shares the homotopy type of the manifold. However, these results focus on smooth manifolds, consider the intrinsic Riemannian metric instead of the Euclidean ambient metric and are not effective since they do not give explicit constants. Nevertheless, they suggest that Rips complexes could be used in practice to produce topologically correct approximations of shapes. If the distances are measured using the  $\ell^\infty$  norm then the Rips complex of  $P$  at scale  $\alpha$  is equal to the Čech complex of  $P$  at scale  $\alpha$  and is also homotopy equivalent to the union of hypercubes of side length  $2\alpha$  centered at the points of  $P$  [14]. In this case the authors in [7] state conditions under which the Rips complex of  $P$  reproduces the homotopy type of the shape sampled by  $P$ . In this paper we suppose distances are measured using the Euclidean norm. Extensions to more general metric spaces will be evoked in the conclusion.

Partially related to our work, we should mention [16] which relates the fundamental group of a Rips complex and its shadow (see below) in dimension 2 and give counterexamples in higher dimensions.

*Sampling conditions.* In any case, it is necessary for a point cloud to be accurate and dense enough to reflect the topology of the shape it samples. The quality of the sample is typically expressed in terms of Hausdorff distance

to the shape. Guaranteed reconstruction methods are generally accompanied by results of the following form: if the Hausdorff distance is smaller than some notion of *topological feature size* of the shape, then the output is topologically correct. First sampling conditions were expressed in terms of the *reach*, which is the infimum of distances between points in the shape and points in its medial axis [4, 2, 11, 34, 19]. Unfortunately, the reach vanishes on sharp concave edges and therefore is not suitable for expressing sampling conditions for non-smooth manifolds or stratified objects. To deal with this problem, authors in [17] introduce a new characterization of the feature size, the  $\mu$ -*reach*, which allows them to formulate sampling conditions for a large class of non-smooth compact subsets of Euclidean space.

In this work, we introduce two new measures of feature size, both called *convexity defects*. Roughly speaking, they measure how far an object is from being locally convex, in the same manner as curvature measures how far an object is from being locally flat. In Section 5, we use these measures to express sampling conditions first for the Čech complex and second for the Rips complex. Regarding the reconstruction with Čech complexes, our work compares well with previous approaches when  $X$  is a smooth set and surprisingly enough, even improve constants when  $X$  has a positive  $\mu$ -reach. Most importantly, this new framework allows us to prove that Rips complexes also provide topologically correct reconstruction, assuming shapes have a positive  $\mu$ -reach, for  $\mu$  sufficiently large.

The remaining sections are organized as follows. In Section 2 we present the necessary background and define the Čech complex and the Rips complex. In Section 3 we introduce and study our two convexity defects functions  $c_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and  $h_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  that we associate with any non-empty bounded subset  $X \subset \mathbb{R}^n$ . Section 4 describes a condition based on  $c_P$  under which the Rips complex of a point set  $P$  at scale  $\alpha$  deformation retracts to the Čech complex of  $P$  at scale  $\alpha$ . This key condition is encapsulated in Theorem 7. In Section 5 we consider a shape  $X$  sampled by a point set  $P$  and formulate conditions under which either Čech complexes or Rips complexes of  $P$  provide approximations of  $X$  with the correct homotopy type. We express those conditions first in terms of an upper bound on  $c_X$  then for shapes with a positive  $\mu$ -reach. Section 6 concludes the paper.

## 2. Background

In this section we introduce the basic definitions and properties needed in the paper and review two well-known examples of simplicial complexes: the Čech complex and the Rips complex.

### 2.1. Metric space, distances, smallest enclosing ball

Throughout this paper, we shall consider subsets of the Euclidean  $n$ -space  $\mathbb{R}^n$  for  $n \geq 1$ . The Euclidean distance between two points  $x$  and  $y$  of  $\mathbb{R}^n$  is denoted  $\|x - y\|$ . Given two subsets  $X$  and  $Y$  of  $\mathbb{R}^n$ , we write  $d_H(Y | X) = \sup_{y \in Y} d(y, X)$  for the *one-sided Hausdorff distance* of  $Y$  from  $X$ , where  $d(y, X)$  is the infimum of the Euclidean distances between  $y$  and points  $x$  in  $X$ . Observe that  $d_H(Y | X) \leq \varepsilon$  if and only if  $Y$  is contained in the  $\varepsilon$ -offset  $X^\varepsilon = \{y \in \mathbb{R}^n \mid d(y, X) \leq \varepsilon\}$ . The *Hausdorff distance* between  $X$  and  $Y$  is  $d_H(X, Y) = \max\{d_H(X | Y), d_H(Y | X)\}$ . Recall that the *diameter* of a subset  $\sigma$  of  $\mathbb{R}^n$  is the supremum of distances between pairs of points in  $\sigma$ , which we denote as  $\text{Diam}(\sigma) = \sup_{p, q \in \sigma} \|p - q\|$ . A subset  $\sigma$  is said to be *bounded* if its diameter is finite.

The closed ball with center  $z$  and radius  $r$  is denoted  $B(z, r)$ . Balls will always be assumed to be closed, unless stated otherwise. It is well known that the smallest ball enclosing a non-empty bounded set  $\sigma$  of  $\mathbb{R}^n$  is well-defined (see Appendix A for a proof). We denote its center by  $\text{Center}(\sigma)$  and its radius by  $\text{Rad}(\sigma)$ . Writing  $\text{Hull}(X)$  for the convex hull of a subset  $X \subset \mathbb{R}^n$  and  $\overline{X}$  for the closure of  $X$ , it is not hard to check (by contradiction) that

$$\text{Center}(\sigma) \in \overline{\text{Hull}(\sigma)}.$$

The following inequalities will be useful in Section 2.4 for relating Čech and Rips complexes:

$$\frac{1}{2} \text{Diam}(\sigma) \leq \text{Rad}(\sigma) \leq \frac{\vartheta_n}{2} \text{Diam}(\sigma) \quad \text{where } \vartheta_n = \sqrt{\frac{2n}{n+1}}. \quad (1)$$

The right most inequality is also known as Jung's Theorem and a short proof can be found in [22]. In particular, we have  $\text{Rad}(\sigma) < \frac{1}{\sqrt{2}} \text{Diam}(\sigma)$  for all dimensions  $n$ .

## 2.2. Abstract simplicial complexes

Let  $P$  be a finite set of points in  $\mathbb{R}^n$ . We call any non-empty subset  $\sigma \subset P$  an *abstract simplex*. Its *dimension* is one less than its cardinality. A  $i$ -simplex is an abstract simplex of dimension  $i$ . If  $\tau \subset \sigma$  is a non-empty subset, we call  $\tau$  a *face* of  $\sigma$  and  $\sigma$  a *coface* of  $\tau$ . An *abstract simplicial complex*  $K$  is a collection of non-empty abstract simplices that contains, with every simplex, the faces of that simplex. The vertex set of the abstract simplicial complex  $K$  is the union of its elements,  $\text{Vert}(K) = \bigcup_{\sigma \in K} \sigma$ . A subcomplex of  $K$  is a simplicial complex  $L \subset K$ . A particular subcomplex is the  $i$ -skeleton consisting of all simplices of dimension  $i$  or less, which we denote by  $K^{(i)}$ . The *shadow* of  $K$  is the subset of  $\mathbb{R}^n$  covered by the convex hull of simplices in  $K$ ,  $\text{Shd } K = \bigcup_{\sigma \in K} \text{Hull}(\sigma)$ , not to be confused with  $|K|$ , the underlying space of a geometric realization of  $K$ ; see [33]. Let  $N$  be the cardinality of the vertex set of  $K$ . The underlying space  $|K|$  of  $K$  can be defined (up to a homeomorphism) by considering a map  $f : \text{Vert}(K) \rightarrow \mathbb{R}^{N-1}$  that sends the  $N$  vertices of  $K$  to  $N$  affinely independent points in  $\mathbb{R}^{N-1}$  and by setting  $|K| = \bigcup_{\sigma \in K} \text{Hull}(f(\sigma))$ . Generally,  $|K|$  and  $\text{Shd } K$  are not homeomorphic, as illustrated in Figure 1.

We now review two natural ways of constructing an abstract simplicial complex, given as input a finite set of points in  $\mathbb{R}^n$  and a feature scale parameter  $t \geq 0$ . The definitions given below appear in different forms in the literature.

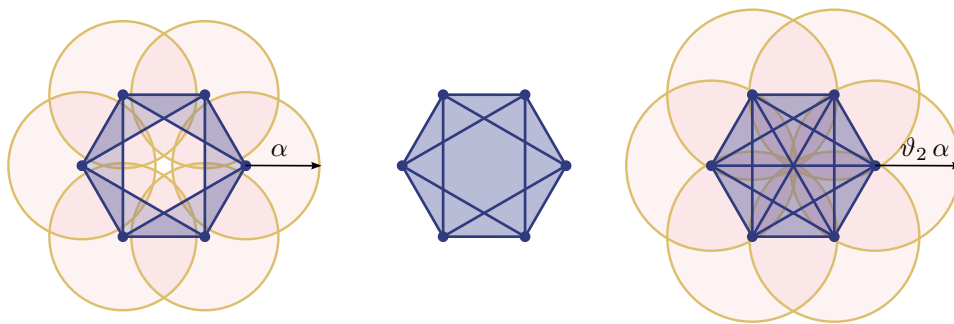


Figure 1: Left: the Čech complex with parameter  $\alpha$  comprises six triangles and is homotopy equivalent to a circle. Middle: the Rips complex with parameter  $\alpha$  contains two more triangles and is homeomorphic to a 2-sphere. Its shadow is a topological disk. Right: the Čech complex with parameter  $\vartheta_2 \alpha$  contains all faces of the 5-simplex and is homeomorphic to a 5-ball.

### 2.3. The Čech complex

The Čech complex  $\mathcal{C}(P, t)$  is the abstract simplicial complex whose  $k$ -simplices correspond to subsets of  $k + 1$  points that can be enclosed in a ball of radius  $t$ ,

$$\mathcal{C}(P, t) = \{\sigma \mid \emptyset \neq \sigma \subset P, \text{Rad}(\sigma) \leq t\}.$$

Equivalently, a  $k$ -simplex  $\{p_0, \dots, p_k\}$  belongs to the Čech complex if and only if the  $k + 1$  closed Euclidean balls  $B(p_i, t)$  have non-empty common intersection. Let  $\text{Nrv } F = \{G \subset F \mid \bigcap G \neq \emptyset\}$  denote the *nerve* of the collection  $F$ . The Čech complex is the nerve of the collection of balls  $\{B(p, t) \mid p \in P\}$ . Since balls are convex, the Nerve Lemma [10, 27] implies that the Čech complex  $\mathcal{C}(P, t)$  is homotopy equivalent to the union of these balls, that is,  $|\mathcal{C}(P, t)| \simeq P^t = \bigcup_{p \in P} B(p, t)$ .

### 2.4. The Rips complex

The *Vietoris-Rips complex* is a variant of the Čech complex which is easier to compute. The Vietoris-Rips complex,  $\mathcal{R}(P, t)$  is the abstract simplicial complex whose  $k$ -simplices correspond to subsets of  $k + 1$  points in  $P$  with diameter at most  $2t$ ,

$$\mathcal{R}(P, t) = \{\sigma \mid \emptyset \neq \sigma \subset P, \text{Diam}(\sigma) \leq 2t\}.$$

For simplicity, we refer to  $\mathcal{R}(P, t)$  as the Rips complex. Recall that the *flag complex* of a graph  $G$ , denoted  $\text{Flag } G$ , is the maximal simplicial complex whose 1-skeleton is  $G$ . The Rips complex is an example of a flag complex. More precisely, this is the largest simplicial complex sharing with the Čech complex the same 1-skeleton,  $\mathcal{R}(P, t) = \text{Flag}(\mathcal{C}(P, t)^{(1)})$ . Generally,  $\mathcal{R}(P, t)$  and  $\mathcal{C}(P, t)$  do not share the same topology; see Figure 1. It follows that the Rips complex  $\mathcal{R}(P, t)$  is generally not homotopy equivalent to the  $t$ -offset  $P^t$ . Our goal in the next section is to find a condition on the point set  $P$  which guarantees that  $|\mathcal{R}(P, t)| \simeq |\mathcal{C}(P, t)|$  and therefore  $|\mathcal{R}(P, t)| \simeq P^t$ . Along the way, we will need a result in [22] which is a consequence of Equation (1) and which says that there is chain of inclusion

$$\mathcal{C}(P, t) \subset \mathcal{R}(P, t) \subset \mathcal{C}(P, \vartheta_n t) \quad \text{where } \vartheta_n = \sqrt{\frac{2n}{n+1}}. \quad (2)$$

### 3. Convexity defects measures

In this section, we introduce and study two functions that one can associate with any non-empty bounded subset  $X \subset \mathbb{R}^n$  and that provide two different ways of measuring convexity defects of  $X$ . Based on the first function, we will formulate in Section 4 a condition which suffices to guarantee that the Rips complex of a finite set of points  $P$  at scale  $\alpha$  deformation retracts to the Čech complex of  $P$  at scale  $\alpha$ . Based on the second function, we will formulate in Section 5 sampling conditions under which the Čech and Rips complexes of a point set  $P$  provide topologically correct reconstructions of a shape  $X$  sampled by the points in  $P$ .

#### 3.1. Definitions and basic properties

To avoid lengthy sentences, we adopt the convention that the subset  $X \subset \mathbb{R}^n$  is always assumed to be non-empty and bounded in this section. In particular, any non-empty subset  $\sigma \subset X$  is also bounded and thus has a well-defined smallest enclosing ball. We first define the set of centers of  $X$  at scale  $t$  as the subset (see Figure 2, left):

$$\text{Centers}(X, t) = \bigcup_{\substack{\emptyset \neq \sigma \subset X \\ \text{Rad}(\sigma) \leq t}} \{\text{Center}(\sigma)\}.$$

Recalling that  $\text{Hull}(X)$  denotes the convex hull of  $X$ , we then extend the notion of convex hull. Specifically, we define the convex hull of  $X$  at scale  $t$  as the subset (see Figure 2, right)

$$\text{Hull}(X, t) = \bigcup_{\substack{\emptyset \neq \sigma \subset X \\ \text{Rad}(\sigma) \leq t}} \text{Hull}(\sigma).$$

If  $X$  is compact then  $\text{Hull}(X, t)$  is a superset of  $\text{Centers}(X, t)$ . If  $P$  is a finite set of points then  $\text{Hull}(P, t)$  is the shadow of the Čech complex  $\mathcal{C}(P, t)$ .

**Definition 1 (Convexity defects functions).** Given a subset  $X \subset \mathbb{R}^n$ , we associate to  $X$  two real-valued functions: the first one  $c_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is defined by  $c_X(t) = d_H(\text{Centers}(X, t) \mid X)$  and the second one  $h_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is defined by  $h_X(t) = d_H(\text{Hull}(X, t) \mid X)$ .



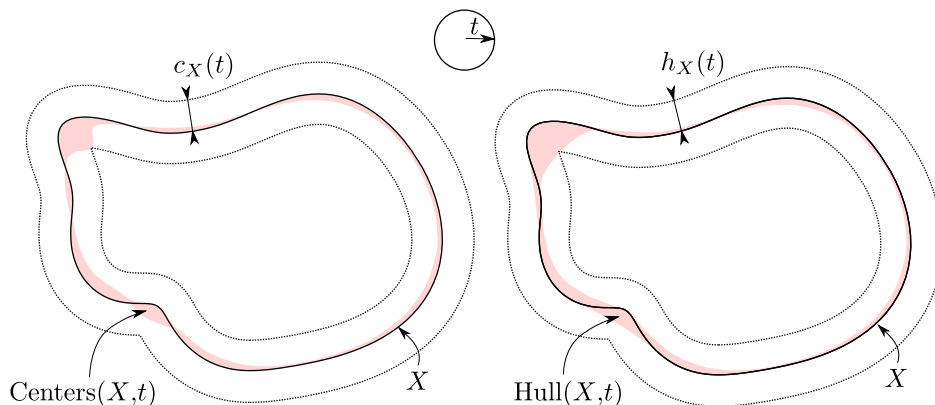


Figure 2: Smallest offset of  $X$  containing  $\text{Centers}(X, t)$  on the left and  $\text{Hull}(X, t)$  on the right.

Intuitively,  $c_X$  and  $h_X$  can be thought of as functions that measure the convexity defects of  $X$  at a given scale. To make this idea precise, observe that if  $X \subset \mathbb{R}^n$  is compact, then we have the three equivalences:  $X$  convex if and only if  $c_X = 0$  if and only if  $h_X = 0$ . The two convexity functions  $c_X$  and  $h_X$  will play a different role. While  $c_P$  is all we need to study the Rips complex of a finite point set  $P$  in Section 4, it turns out that  $h_X$  is more stable than  $c_X$  and will be used in Section 5.1 to express sampling conditions in reconstruction theorems. We plotted the graph of the function  $c_P$  for various finite point sets in Figure 10.

Before studying in more details functions  $c_X$  and  $h_X$  in the next two sections, let us make some brief remarks. Because  $X$  is a subset of both  $\text{Centers}(X, t)$  and  $\text{Hull}(X, t)$ , it follows that the two one-sided Hausdorff distances  $d_H(X \mid \text{Centers}(X, t))$  and  $d_H(X \mid \text{Hull}(X, t))$  vanish. Hence, we could have used in the above definition two-sided Hausdorff distances instead of one-sided Hausdorff distances. The two functions  $c_X$  and  $h_X$  both vanish at 0, are increasing in the interval  $[0, \text{Rad}(X)]$  and become constant above  $\text{Rad}(X)$ . Since  $\text{Center}(\sigma) \in \overline{\text{Hull}(\sigma)}$ , we have  $c_X \leq h_X$ . It is easy to check that for a subset  $X \subset \mathbb{R}^n$  and two non-negative real numbers  $t$  and  $\alpha$ , the following three conditions are equivalent: (1)  $h_X(t) \leq \alpha$ ; (2)  $\text{Hull}(X, t) \subset X^\alpha$ ; (3)  $[\text{Rad}(\sigma) \leq t \implies \text{Hull}(\sigma) \subset X^\alpha]$  for all  $\sigma \subset X$ . In particular, we get that  $h_X(t) \leq t$  for all  $t \geq 0$  since  $\text{Rad}(\sigma) \leq t \implies \text{Hull}(\sigma) \subset \sigma^t$  as a direct consequence of Lemma 1 (below) applied for  $x = y$ .

**Lemma 1.** *For any non-empty bounded subset  $\sigma \subset \mathbb{R}^n$ , any point  $x \in \mathbb{R}^n$*

and any point  $y \in \text{Hull}(\sigma)$ , we have that  $d(x, \sigma)^2 \leq \|x - y\|^2 + \text{Rad}(\sigma)^2 - \|y - \text{Center}(\sigma)\|^2$ .

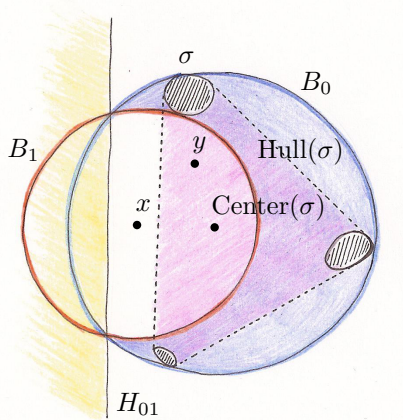


Figure 3: Notation for the proof of Lemma 1.

*Proof.* Suppose  $d(x, \sigma) > \|x - y\|$  for otherwise the result is clear. Let  $B_0$  be the smallest ball enclosing  $\sigma$  and let  $B_1$  be the largest ball centered at  $x$  whose interior does not intersect  $\sigma$ ; see Figure 3. By construction,  $\sigma \subset B_0 \setminus B_1$ . Recall that the power distance of a point  $y$  from a ball  $B$  is  $\pi_B(y) = \|y - z\|^2 - r^2$ , where  $z$  is the center of  $B$  and  $r$  its radius. Let  $H_{01}$  be the set of points whose power distance to  $B_0$  is at most as large as the power distance to  $B_1$ .  $H_{01}$  is a closed half-space which contains the set difference  $B_0 \setminus B_1$ . In particular, it contains  $\sigma$  and any point  $y \in \text{Hull}(\sigma)$ . Thus,  $\pi_{B_0}(y) \leq \pi_{B_1}(y)$  and the result follows.  $\square$

### 3.2. Characterizing critical values of the distance function

In the previous section we noted that  $c_X(t) \leq h_X(t) \leq t$  for all  $t$ . The goal of this section is to establish that equality is attained if and only if  $t$  is a critical value of the distance function to  $X$ . This property will not be used before Section 5 but sheds light on results of Section 4.

We need some definitions. The distance function  $d(\cdot, X)$  to the compact set  $X \subset \mathbb{R}^n$  maps every point  $y \in \mathbb{R}^n$  to its Euclidean distance to  $X$ ,  $d(y, X) = \min_{x \in X} \|x - y\|$ . Although the distance function is not differentiable, it is possible to define a notion of critical points analogue to the

classical one for differentiable functions as illustrated in Figure 4. Specifically, Grove defines in [29, page 360] critical points for the distance function to a closed subset of a Riemannian manifold. We recast this definition in our context as follows. Let  $\Gamma_X(y) = \{x \in X \mid d(y, X) = \|x - y\|\}$  be the set of points in  $X$  closest to  $y$ :

**Definition 2.** We say that  $y \in \mathbb{R}^n$  is a *critical point* of the distance function  $d(\cdot, X)$  if  $y \in \text{Hull}(\Gamma_X(y))$ . The *critical values* of  $d(\cdot, X)$  are the images by  $d(\cdot, X)$  of its critical points.

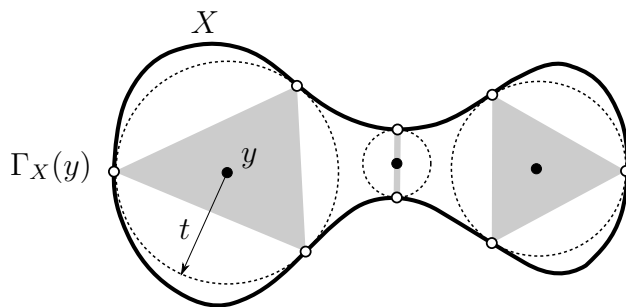


Figure 4: The black points are the critical points of the curve  $X$ .

Next lemma provides two characterizations of the critical values of the distance function to a compact set  $X \subset \mathbb{R}^n$ , based respectively on the two convexity defects functions  $c_X$  and  $h_X$ .

**Lemma 2.** For any compact set  $X \subset \mathbb{R}^n$  and any real number  $t > 0$ , the following three conditions are equivalent: (1)  $t$  is a critical value of  $d(\cdot, X)$ ; (2)  $c_X(t) = t$ ; (3)  $h_X(t) = t$ .

*Proof.* Consider a non-empty bounded subset  $\sigma \subset \mathbb{R}^n$  and a point  $y \in \mathbb{R}^n$ . Making  $x = y$  in Lemma 1, we observe that if  $y \in \text{Hull}(\sigma)$  satisfies  $d(y, \sigma) \geq t$  and  $\text{Rad}(\sigma) \leq t$ , then  $y = \text{Center}(\sigma)$  and  $t = \text{Rad}(\sigma)$ .

Let us prove that (1)  $\implies$  (2). Consider a critical point  $y$  whose distance to  $X$  is  $t$  and set  $\sigma = \Gamma_X(y)$  as shown in Figure 4. By definition  $y \in \text{Hull}(\sigma)$  and by construction  $d(y, \sigma) = t$  and  $\text{Rad}(\sigma) \leq t$ . Thanks to our observation, it follows that  $y = \text{Center}(\sigma)$  and consequently  $c_X(t) = t$ . The implication (2)  $\implies$  (3) follows from  $c_X(t) \leq h_X(t) \leq t$ . Let us prove that (3)  $\implies$  (1). In other words, suppose  $h_X(t) = t$  and let us prove that  $t$  is a critical value of  $d(\cdot, X)$ . Since  $X$  is compact,  $h_X(t) = t$  means that we can find

a compact set  $\emptyset \neq \sigma \subset X$  with  $\text{Rad}(\sigma) \leq t$  and  $y \in \text{Hull}(\sigma)$  such that  $t = d(y, X) \leq d(y, \sigma)$ . Our observation then implies that  $y = \text{Center}(\sigma)$ ,  $t = \text{Rad}(\sigma)$  and  $\sigma$  represents a set of points in  $X$  with minimum distance to  $y$ . Since  $y \in \text{Hull}(\sigma) \subset \text{Hull}(\Gamma_X(y))$ , it follows that  $y$  is a critical point of the distance function to  $X$ , which concludes the proof.  $\square$

An adaptation of Morse theory to distance functions tells us that changes in the topology of  $t$ -offsets  $X^t$  occur when the offset parameter  $t$  reaches critical values of the distance function to  $X$ . Indeed, slightly recasting Proposition 1.8 in [29, page 362], we have:

**Theorem 3** (Isotopy Theorem [29]). *Let  $X \subset \mathbb{R}^n$  be a compact set and let  $\beta \geq \alpha > 0$  be two real numbers. If the distance function  $d(\cdot, X)$  has no critical values in the interval  $[\alpha, \beta]$ , then  $X^\beta$  deformation retracts to  $X^\alpha$ .*

Our characterizations of critical values allow us to reexpress the condition in the above theorem. To be specific, we get that if  $c_X(t) < t$  for all  $t \in [\alpha, \beta]$ , then  $X^\beta$  deformation retracts to  $X^\alpha$ . Replacing  $X$  by a finite point set  $P$  and using that the Čech complex  $\mathcal{C}(P, t)$  is homotopy equivalent to the  $t$ -offset  $P^t$ , we obtain that  $\mathcal{C}(P, \beta)$  is homotopy equivalent to  $\mathcal{C}(P, \alpha)$  whenever  $c_P(t) < t$  for all  $t \in [\alpha, \beta]$ . We shall see in Section 4 that under the same condition a stronger result holds, namely the existence of a sequence of collapses from  $\mathcal{C}(P, \beta)$  to  $\mathcal{C}(P, \alpha)$ . Strengthening this condition, we will be able to guarantee the existence of a sequence of collapses from  $\mathcal{R}(P, \beta)$  to  $\mathcal{C}(P, \alpha)$ . Variants of this condition will then be devised in Section 5 to ensure topologically correct reconstruction of shapes by Čech or Rips complexes.

### 3.3. Stability

In this section, we state the stability of  $c_X$  and  $h_X$ . For technical reasons, we need the stability of  $c_P$  under perturbations of  $P$  at the end of the proof of Theorem 7 in Section 4 to relax the assumption that the finite point set  $P$  that we consider is in general position. The stability of  $h_X$  under perturbation of  $X$  will be crucial for establishing reconstruction theorems in Section 5.

**Lemma 4.** *For every pair of subsets  $X$  and  $Y$  of  $\mathbb{R}^n$  such that  $d_H(X, Y) \leq \varepsilon$  and for every  $t \geq 0$ , we have*

$$c_Y(t) \leq c_X(t + \varepsilon) + \sqrt{2t\varepsilon + \varepsilon^2} + \varepsilon.$$

*Proof.* Consider a non-empty subset  $\sigma \subset Y$  with  $\text{Rad}(\sigma) \leq t$  and set  $\xi = X \cap \sigma^\varepsilon$ . By construction,  $\xi$  is non-empty and  $d_H(\xi, \sigma) \leq \varepsilon$ . Hence, setting  $\delta = \sqrt{2t\varepsilon + \varepsilon^2}$ , Lemma 16 implies that  $\text{Rad}(\xi) \leq t + \varepsilon$  and  $\|\text{Center}(\sigma) - \text{Center}(\xi)\| \leq \delta$ . We get

$$\text{Center}(\sigma) \subset \text{Center}(\xi)^\delta \subset X^{c_X(t+\varepsilon)+\delta} \subset Y^{c_X(t+\varepsilon)+\delta+\varepsilon},$$

yielding to the result.  $\square$

**Lemma 5.** *For every pair of subsets  $X$  and  $P$  of  $\mathbb{R}^n$  such that  $d_H(X, P) \leq \varepsilon$  and for every  $t \geq 0$ , we have  $h_P(t) \leq h_X(t + \varepsilon) + 2\varepsilon$ .*

*Proof.* Consider a non-empty subset  $\sigma \subset P$  with  $\text{Rad}(\sigma) \leq t$  and set  $\xi = X \cap \sigma^\varepsilon$ . By construction,  $\xi$  is non-empty and  $d_H(\xi, \sigma) \leq \varepsilon$ . Hence, Lemma 16 implies that  $\text{Rad}(\xi) \leq t + \varepsilon$ . Using  $\text{Hull}(\xi^\varepsilon) = \text{Hull}(\xi)^\varepsilon$ , we get that  $\text{Hull}(\sigma) \subset \text{Hull}(\xi)^\varepsilon \subset X^{h_X(t+\varepsilon)+\varepsilon} \subset P^{h_X(t+\varepsilon)+2\varepsilon}$ , yielding the result.  $\square$

#### 4. From Rips to Čech complexes

In this section, we introduce a 2-parameter family of Rips complexes and give the precise condition on a finite point set for which we can prove that a Rips complex in this family deformation retracts to a Čech complex. We begin by defining this family of Rips complexes and state our results in Section 4.1. We then introduce the tools we need to prove our results in Section 4.2. The proofs are presented in Section 4.3.

##### 4.1. Quasi Rips complexes and statement of results

Following [16], we first define a 2-parameter family that contains prior Rips complexes as a subfamily. The motivation for this construction is to account for the uncertainty of measures by allowing uncertainty of the edges belonging to Rips complexes in the family; see [16].

**Definition 3.** For any point set  $P \subset \mathbb{R}^n$  and any real numbers  $\alpha, \beta \geq 0$  with  $\alpha \leq \beta$ , we call the flag complex of any graph  $G$  satisfying  $\mathcal{R}(P, \alpha) \subset \text{Flag } G \subset \mathcal{R}(P, \beta)$  an  $(\alpha, \beta)$ -quasi Rips complex of  $P$ .

In other words, the simplicial complex  $\text{Flag } G$  is an  $(\alpha, \beta)$ -quasi Rips complex of  $P$  if and only if every pairs of points in  $P$  within distance  $2\alpha$  are connected by an edge in  $G$  and no edge of  $G$  has length larger than  $2\beta$ .

Equivalently, for every pairs  $(p, q) \in P^2$ ,  $\|p - q\| \leq 2\alpha$  implies  $pq \in G$  and  $\|p - q\| > 2\beta$  implies  $pq \notin G$ . In particular,  $K$  is an  $(\alpha, \alpha)$ -quasi Rips complex of  $P$  if and only if  $K = \mathcal{R}(P, \alpha)$ .

To state our results, it is convenient to define  $\alpha$  to be an *inert value* of  $P$  if  $\text{Rad}(\sigma) \neq \alpha$  for all non-empty subsets  $\sigma \subset P$ . The finiteness of  $P$  implies that  $P$  has only finitely many non-inert values. Thus, assuming  $\alpha$  to be inert is not a too restrictive hypothesis.

As a warm-up, we first state conditions in Theorem 6 (below) under which there exists a sequence of elementary collapses which transform one Čech complex into another one. We recall that an *elementary collapse* is the operation that removes a pair of simplices  $(\sigma, \tau)$  from a simplicial complex  $K$  assuming  $\tau$  is the unique proper coface of  $\sigma$  in  $K$ . The result is a simplicial complex  $K \setminus \{\sigma, \tau\}$  to which  $K$  deformation retracts.

**Theorem 6.** *Let  $P \subset \mathbb{R}^n$  be a finite set of points. For any real numbers  $\beta \geq \alpha \geq 0$  such that  $\alpha$  is an inert value of  $P$  and  $c_P(t) < t$  for all  $t \in [\alpha, \beta]$ , there exists a sequence of elementary collapses from  $\mathcal{C}(P, \beta)$  to  $\mathcal{C}(P, \alpha)$ .*

Theorem 6 can be thought of as a combinatorial version of the Isotopy theorem presented in Section 3.2. We are now ready to state our main theorem:

**Theorem 7.** *Let  $P \subset \mathbb{R}^n$  be a finite set of points. For any real numbers  $\beta \geq \alpha \geq 0$  such that  $\alpha$  is an inert value of  $P$  and  $c_P(\vartheta_n \beta) < 2\alpha - \vartheta_n \beta$ , there exists a sequence of elementary collapses from any  $(\alpha, \beta)$ -quasi Rips complex of  $P$  to the Čech complex  $\mathcal{C}(P, \alpha)$ .*

Note that choosing  $\beta = \alpha$  in the theorem gives conditions under which  $|\mathcal{R}(P, \alpha)| \simeq |\mathcal{C}(P, \alpha)| \simeq P^\alpha$ . Figure 7 on the left provides a graphical representation of the hypothesis of the theorem. Let us sketch quickly the proof of Theorem 7. Observe that the condition  $c_P(\vartheta_n \beta) < 2\alpha - \vartheta_n \beta$  implies that  $c_P(t) < t$  for all  $t \in [\alpha, \vartheta_n \beta]$  since

$$c_P(t) \leq c_P(\vartheta_n \beta) < 2\alpha - \vartheta_n \beta \leq \alpha \leq t,$$

whenever  $\alpha \leq t \leq \vartheta_n \beta$ . Theorem 6 then implies that there exists a sequence of collapses reducing  $\mathcal{C}(P, \vartheta_n \beta)$  to  $\mathcal{C}(P, \alpha)$ . Since any  $(\alpha, \beta)$ -quasi Rips complex  $\text{Flag } G$  is nested between  $\mathcal{C}(P, \alpha)$  and  $\mathcal{C}(P, \vartheta_n \beta)$ , the key idea in the proof of the above theorem is to monitor changes in the complex  $\mathcal{C}(P, t) \cap \text{Flag } G$  as we decrease the scale parameter  $t$  from  $\vartheta_n \beta$  to  $\alpha$ , that is as we go from  $\mathcal{C}(P, \vartheta_n \beta)$  to  $\text{Flag } G$ . The proof is given in Section 4.3. Before embarking in the proof, we will define collapses and extended collapses in the next section.

#### 4.2. Extended collapses

In this section, we introduce *collapses* and *extended collapses* which will turn out to be convenient to describe changes that occur in the families of complexes that we consider in our proof of Theorem 7.

We need some definitions. The inclusion defines a partial order relation on simplices. Given a set of simplices  $\Sigma$  and a simplex  $\sigma \in \Sigma$ , we say that  $\sigma$  is *inclusion-maximal* in  $\Sigma$  if  $\sigma$  has no proper coface in  $\Sigma$ . Similarly, we say that  $\sigma$  is *inclusion-minimal* if it has no proper face in  $\Sigma$ . When clear from the context we will omit  $\Sigma$ . Suppose  $\sigma$  is a simplex of the simplicial complex  $K$ . The *star* of  $\sigma$  in  $K$ , denoted  $\text{St}_K(\sigma)$ , is the collection of simplices of  $K$  having  $\sigma$  as a face. The closure of  $\text{St}_K(\sigma)$  is denoted  $\overline{\text{St}}_K(\sigma)$ ; it is the smallest simplicial complex containing  $\text{St}_K(\sigma)$ . The *link* of  $\sigma$  in  $K$ , denoted  $\text{Lk}_K(\sigma)$ , is the collection of simplices of  $K$  lying in  $\overline{\text{St}}_K(\sigma)$  that are disjoint from  $\sigma$ . A simplicial complex  $K$  is said to be a *cone* if it contains a vertex  $o$  such that the following implication holds:  $\sigma \in K \implies \sigma \cup \{o\} \in K$ . The vertex  $o$  is called the *apex* of the cone. By definition a cone can never be empty since it always contains at least its apex.

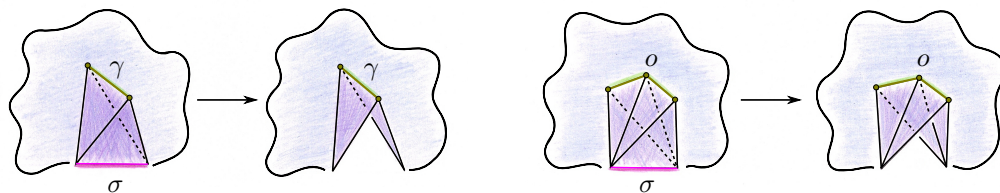


Figure 5: Left: In a classical collapse, the link of  $\sigma$  has a unique inclusion-maximal simplex  $\gamma$ . Equivalently, the star of  $\sigma$  has a unique inclusion-maximal simplex  $\sigma \cup \gamma$  different from  $\sigma$ . Right: In an extended collapse, the link of  $\sigma$  is a cone with apex  $o$ .

Given a simplicial complex  $K$ , we are interested in the operation that removes the entire star of a simplex  $\sigma \in K$  as illustrated in Figure 5. Provided that there is a unique inclusion-maximal simplex  $\tau \neq \sigma$  in the star of  $\sigma$ , it is well-known that  $|K|$  deformation retracts to  $|K \setminus \text{St}_K(\sigma)|$  and the operation that removes  $\text{St}_K(\sigma)$  is then called a *collapse* [25]. Any collapse can be decomposed into a finite sequence of elementary collapses. Following and extending what was done in [9], we call the operation that removes  $\text{St}_K(\sigma)$  assuming the weaker condition that the link of  $\sigma$  is a cone an *extended collapse*. Our terminology finds its justification in the following lemma.

**Lemma 8.** *Let  $K$  be a simplicial complex and let  $\sigma$  be a simplex of  $K$ . If the link of  $\sigma$  is a cone, then there is a sequence of collapses from  $K$  to  $K \setminus \text{St}_K(\sigma)$ .*

*Proof.* First, we establish that a cone  $L$  can always be reduced to its apex  $o$  by a sequence of collapses. Suppose  $L \neq \{o\}$  for otherwise the result is clear and consider an inclusion-maximal simplex in  $L$ . It has the form  $\alpha \cup \{o\}$  with  $o \notin \alpha \in L$ . Let us prove that the operation that removes the pair of simplices  $(\alpha, \alpha \cup \{o\})$  is a collapse. If  $\beta$  is a coface of  $\alpha$ , then  $\beta \cup \{o\}$  is a coface of  $\alpha \cup \{o\}$ . Since  $\alpha \cup \{o\}$  is inclusion-maximal, the only possibility is that  $\beta \cup \{o\} = \alpha \cup \{o\}$ , showing that all cofaces of  $\alpha$  are faces of  $\alpha \cup \{o\}$ . Thus, removing the pair  $(\alpha, \alpha \cup \{o\})$  is a collapse and the result is still a cone with apex  $o$  but with a smaller number of simplices. By repeating the process we eventually get a complex reduced to  $\{o\}$ .

Now, suppose the link of  $\sigma$  in  $K$  is a cone  $L$  with apex  $o$ . We deduce from the previous sequence of collapses that reduces  $L$  to  $\{o\}$  a sequence of collapses that reduces  $K$  to  $K \setminus \text{St}_K(\sigma)$  as follows. To each collapse that removes the pair  $(\alpha, \alpha \cup \{o\})$  in  $L$  described above, we associate the collapse that removes the pair  $(\sigma \cup \alpha, \sigma \cup \alpha \cup \{o\})$  in  $K$ . Indeed, since  $\alpha$  has a unique proper coface  $\alpha \cup \{o\}$  in  $L$ , it follows that  $\sigma \cup \alpha$  has a unique proper coface  $\sigma \cup \alpha \cup \{o\}$  in  $K$ . At the end of this sequence of collapses, the link of  $\sigma$  is reduced to  $\{o\}$ . After a last collapse that removes the pair  $(\sigma, \sigma \cup \{o\})$  we get  $K \setminus \text{St}_K(\sigma)$ .  $\square$

### 4.3. Proof of results

We will prove directly Theorem 7 and omit the proof of Theorem 6 since one can easily derive a proof of Theorem 6 by slightly adapting the first part of the proof below.

*Proof of Theorem 7.* Let  $G$  be a graph whose flag complex is an  $(\alpha, \beta)$ -quasi Rips complex of  $P$ . For  $t \geq 0$ , consider the simplicial complex  $\mathcal{F}(t) = \mathcal{C}(P, t) \cap \text{Flag } G$ . Clearly, we have the chain of inclusions:

$$\mathcal{C}(P, \alpha) \subset \mathcal{R}(P, \alpha) \subset \text{Flag } G \subset \mathcal{R}(P, \beta) \subset \mathcal{C}(P, \vartheta_n \beta)$$

and therefore  $\mathcal{F}(\alpha) = \mathcal{C}(P, \alpha)$  and  $\mathcal{F}(\vartheta_n \beta) = \text{Flag } G$ . As we continuously increase the feature parameter  $t$  from  $\alpha$  to  $\vartheta_n \beta$ , we get a finite family of nested Čech complexes:

$$\mathcal{C}(P, \alpha) = \mathcal{C}_0 \subset \mathcal{C}_1 \subset \dots \subset \mathcal{C}_k = \mathcal{C}(P, \vartheta_n \beta).$$



For  $0 < i < k$ , let  $t_i$  be the smallest value of  $t$  such that  $\mathcal{C}_i = \mathcal{C}(P, t)$  and set  $\mathcal{F}_i = \mathcal{F}(t_i)$ . In particular,  $\mathcal{C}_i = \mathcal{C}(P, t_i)$  and  $\mathcal{F}_i = \mathcal{C}_i \cap \text{Flag } G$ . Correspondingly, we get a 1-parameter family of simplicial complexes by intersecting each complex  $\mathcal{C}_i$  in the above sequence with  $\text{Flag } G$ :

$$\mathcal{C}(P, \alpha) = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_k = \text{Flag } G.$$

Let us first assume that  $P$  satisfies the two generic conditions  $(\star)$  and  $(\star\star)$  instead of the condition that  $\text{Rad}(\sigma) \neq \alpha$  for all non-empty subsets  $\sigma \subset P$ :

- $(\star)$  For all simplices  $\sigma, \tau \subset P$ , if  $\text{Rad}(\sigma) = \text{Rad}(\tau)$  then  $\text{Center}(\sigma) = \text{Center}(\tau)$ ;
- $(\star\star)$  For any ball  $B$ , the set of simplices in  $P$  that have  $B$  as a smallest enclosing ball is either empty or has a unique inclusion-minimal element.

Under these two conditions, we prove the theorem in two stages. First, we show that  $\mathcal{C}_{i-1}$  is a collapse of  $\mathcal{C}_i$  for all  $0 < i \leq k$ . Secondly, we show that  $\mathcal{F}_{i-1}$  is either equal or an extended collapse of  $\mathcal{F}_i$  for all  $0 < i \leq k$ .

(a) Because of condition  $(\star)$ , all simplices in the set difference  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$  share the same smallest enclosing ball  $B(z_i, t_i)$  with center  $z_i$  and radius  $t_i$ . Because of condition  $(\star\star)$ , the set of simplices sharing the same smallest enclosing ball  $B(z_i, t_i)$  has a unique inclusion-minimal element  $\sigma_i$ . Our plan is to prove that  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$  is the star of  $\sigma_i$  and has a unique inclusion-maximal element  $\tau_i \neq \sigma_i$  which will entail that  $\mathcal{C}_i$  collapses to  $\mathcal{C}_{i-1}$ . Suppose  $\eta$  is a coface of  $\sigma_i$  in  $\mathcal{C}_i$ . Since  $\sigma_i \subset \eta$ , we deduce that  $t_i = \text{Rad}(\sigma_i) \leq \text{Rad}(\eta)$  and therefore  $\eta \in \mathcal{C}_i \setminus \mathcal{C}_{i-1}$ . Hence,  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$  is the star of  $\sigma_i$  in  $\mathcal{C}_i$ . Note that the simplex  $\tau_i = \{p \in P \mid \|z_i - p\| \leq t_i\}$  obtained by gathering all points of  $P$  in  $B(z_i, t_i)$  belongs to  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$  and is the unique inclusion-maximal simplex in this set difference (see Figure 6, left). To prove that  $\mathcal{C}_i$  collapses to  $\mathcal{C}_{i-1}$ , it remains to establish that  $\sigma_i \neq \tau_i$ . By choice of  $\sigma_i$  as an inclusion-minimal element amongst simplices with smallest enclosing ball  $B(z_i, t_i)$ , the vertices of  $\sigma_i$  all lie on the sphere with center  $z_i$  and radius  $t_i$ . On the other hand, by definition of  $c_P(t_i)$  as the one-sided Hausdorff distance of the centers of  $P$  at scale  $t_i$  from  $P$ , there exists at least a point  $o$  of  $P$  at distance  $c_P(t_i)$  or less from the center  $z_i$ . Since  $c_P(t_i) \leq c_P(\vartheta_n \beta) < \alpha \leq t_i$ , the point  $o$  belongs to the interior of  $B(z_i, t_i)$ . Thus,  $o \notin \sigma_i$ ,  $o \in \tau_i$ , and therefore  $\sigma_i \neq \tau_i$ , showing that  $\mathcal{C}_i$  collapses to  $\mathcal{C}_{i-1}$ .

(b) Let us now turn our attention to  $\mathcal{F}_i$  and  $\mathcal{F}_{i-1}$ . If  $\sigma_i \notin \mathcal{F}_i$ , then  $\mathcal{F}_i = \mathcal{F}_{i-1}$ . If  $\sigma_i \in \mathcal{F}_i$ , the star of  $\sigma_i$  in  $\mathcal{F}_i$  is equal to the star of  $\sigma_i$  in  $\mathcal{C}_i$

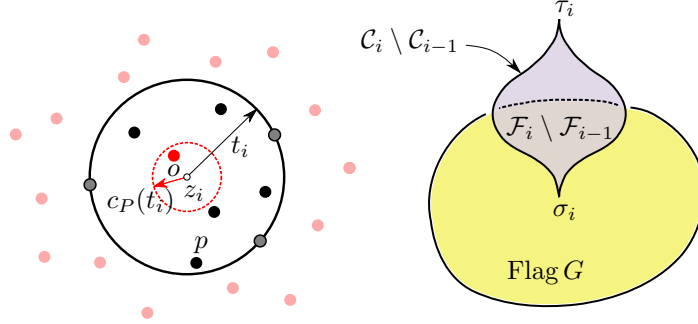


Figure 6: Notation for the proof of Theorem 7. Left:  $\tau_i$  is the simplex whose vertices are points of  $P$  in  $B(z_i, t_i)$ .  $\sigma_i$  is the face obtained by keeping vertices on the boundary of  $B(z_i, t_i)$ . Right: Schematic representation of simplices in  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$ .

intersection the flag of  $G$  and  $\mathcal{F}_{i-1} = \mathcal{F}_i \setminus \text{St}_{\mathcal{F}_i}(\sigma_i)$  (see Figure 6, right). Let us prove that the link of  $\sigma_i$  in  $\mathcal{F}_i$  is a cone with apex  $o$ , which guarantees that  $\mathcal{F}_{i-1}$  is an extended collapse of  $\mathcal{F}_i$ . Suppose  $\eta$  is a coface of  $\sigma_i$  in  $\mathcal{F}_i$  and let us show that  $\eta \cup \{o\}$  is also a coface. Clearly,  $\eta \cup \{o\}$  belongs to the Čech complex  $\mathcal{C}_i$  since for all points  $p \in \eta \cup \{o\}$ ,  $\|z_i - p\| \leq t_i$ . Let us prove that  $\eta \cup \{o\}$  also belongs to  $\text{Flag } G$ . Since  $\eta$  belongs to  $\text{Flag } G$ , it suffices to prove that all edges connecting  $o$  to a vertex  $p$  of  $\eta$  have length  $2\alpha$  or less. Indeed, for all points  $p \in \eta$ , we have

$$\|p - o\| \leq \|z_i - p\| + \|z_i - o\| \leq t_i + c_P(t_i) \leq 2\alpha$$

showing that  $\eta \cup \{o\} \in \text{Flag } G$ . Hence,  $\eta \cup \{o\}$  belongs to  $\mathcal{F}_i$ . Setting  $\eta = \sigma_i$ , we get that  $\sigma_i \cup \{o\}$  is a coface of  $\sigma_i$  and since  $o \notin \sigma_i$ , it follows that  $\{o\}$  belongs to the link of  $\sigma_i$  in  $\mathcal{F}_i$ . Hence, the link of  $\sigma_i$  in  $\mathcal{F}_i$  is a cone, which concludes the proof of Theorem 7 assuming generic conditions  $(\star)$  and  $(\star\star)$  instead of the condition  $\text{Rad}(\sigma) \neq \alpha$  for all non-empty subsets  $\sigma \subset P$ .

If  $P$  does not satisfy the generic conditions  $(\star)$  and  $(\star\star)$ , we use Lemma 17 in Appendix B to find a perturbation  $f$  of the points such that  $f(P)$  satisfies  $(\star)$  and  $(\star\star)$  and conditions (i), (ii) and (iii) of Lemma 17 for some  $\beta' > \beta$ . Applying Theorem 7 to  $f(P)$  with the values  $\alpha$  and  $\beta'$ , we get that there exists a sequence of collapses from the  $(\alpha, \beta')$ -quasi Rips complex  $\text{Flag } f(G) = f(\text{Flag}(G))$  to the Čech complex  $\mathcal{C}(f(P), \alpha) = f(\mathcal{C}(P, \alpha))$ . Hence, the theorem also holds in the non-generic case.  $\square$

## 5. Shape reconstruction

In this section, we are interested in reconstructing a compact set  $X \subset \mathbb{R}^n$  only known through a finite set of possibly noisy points  $P \subset \mathbb{R}^n$ . Using the convexity defect function  $h_X$ , we formulate two sampling conditions which guarantee respectively that the Čech complex and the Rips complex of  $P$  are homotopy equivalent to any arbitrarily small offset of  $X$  (Section 5.1). We then construct a bridge between shapes with an upper bounded convexity defects function and shapes with a lower bounded critical function in Section 5.2. Finally, we compute in Section 5.3 the lowest density of points authorized by our theorems for a correct reconstruction of shapes with a positive  $\mu$ -reach.

### 5.1. Sampling conditions based on convexity defects functions

We assemble the pieces and deduce conditions under which the Čech complex and the Rips complex of a finite set of points retrieve the topology of the shape the points sample. Throughout the section,  $X$  designates a compact subset of  $\mathbb{R}^n$  and  $P$  is a finite set of points, whose Hausdorff distance to  $X$  is  $\varepsilon$  or less.

*Reconstruction with the Čech complex.* The assumption that  $d_H(X, P) \leq \varepsilon$  implies the following chain of inclusions:

$$P^\alpha \subset X^{\alpha+\varepsilon} \subset P^{\alpha+2\varepsilon} \subset X^{\alpha+3\varepsilon}.$$

From [7], we know that whenever we consider four nested spaces  $P_0 \subset X_0 \subset P_1 \subset X_1$  such that  $X_1$  deformation retracts to  $X_0$  and  $P_1$  deformation retracts to  $P_0$ , then  $X_0$  deformation retracts to  $P_0$ . Applying this result to our context combined with the Isotopy Theorem and the characterization of critical points given in Lemma 2, we deduce immediately that  $X^{\alpha+\varepsilon}$  deformation retracts to  $P^\alpha$  whenever the following two conditions are fulfilled:

$$\begin{aligned} h_X(t) &< t, & \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon], \\ h_P(t) &< t, & \forall t \in [\alpha, \alpha + 2\varepsilon]. \end{aligned}$$

Since  $d_H(X, P) \leq \varepsilon$ , Lemma 5 implies that  $h_P(t) \leq h_X(t + \varepsilon) + 2\varepsilon$  and therefore the above two conditions are fulfilled as soon as the following stronger condition holds:  $h_X(t) < t - 3\varepsilon$ ,  $\forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$ . Because  $h_X$  is non-negative, this condition implies that  $2\varepsilon < \alpha$ . Because  $h_X$  is increasing, it also implies that  $h_X(t) < t$  for all  $t \in [\alpha - 2\varepsilon, \alpha + 3\varepsilon]$ , showing that  $\eta$ -offsets of  $X$  for  $\eta$  in the interval  $[\alpha - 2\varepsilon, \alpha + 3\varepsilon]$  are all homotopy equivalent. We summarize our findings in the following theorem:

**Theorem 9.** *Let  $\varepsilon, \alpha > 0$  such that  $2\varepsilon < \alpha$ . Let  $P$  be a finite set of points whose Hausdorff distance to a compact subset  $X$  is  $\varepsilon$  or less. The Čech complex  $\mathcal{C}(P, \alpha)$  is homotopy equivalent to  $X^\eta$  for all  $\eta \in [\alpha - 2\varepsilon, \alpha + 3\varepsilon]$  whenever  $h_X(t) < t - 3\varepsilon$  for all  $t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$ .*

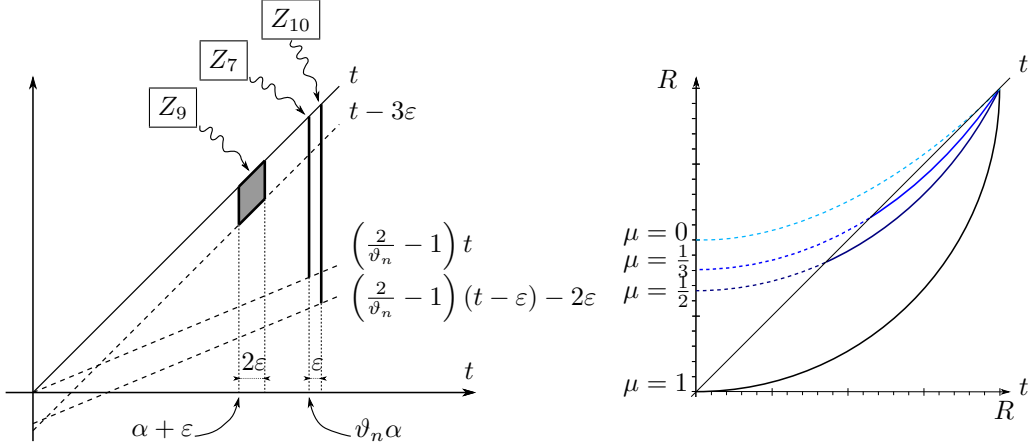


Figure 7: Left: For  $i \in \{7, 9, 10\}$ , the hypotheses of Theorem  $i$  are depicted as regions  $Z_i$  avoided by the graph of a convexity defects function. Specifically, if  $c_P \cap Z_7 = \emptyset$ , Theorem 7 implies  $\mathcal{R}(P, \alpha) \simeq P^\alpha$ . If  $h_X \cap Z_9 = \emptyset$ , Theorem 9 implies  $\mathcal{C}(P, \alpha) \simeq X^{\alpha - 2\varepsilon}$ . If  $h_X \cap Z_{10} = \emptyset$ , Theorem 10 implies  $\mathcal{R}(P, \alpha) \simeq X^{2\alpha - \vartheta_n \alpha - 2\varepsilon}$ . Right: Upper bound on  $h_X$  for  $\mu \in \{0, \frac{1}{3}, \frac{1}{2}, 1\}$  provided by Lemma 12.

*Reconstruction with the Rips complex.* If furthermore we suppose that the condition  $c_P(\vartheta_n \beta) < 2\alpha - \vartheta_n \beta$  holds, we can apply Theorem 7 and deduce that  $(\alpha, \beta)$ -quasi Rips complexes of  $P$  deformation retracts to the Čech complex  $\mathcal{C}(P, \alpha)$ . Using Lemma 5, we get that  $c_P(\vartheta_n \beta) \leq h_P(\vartheta_n \beta) \leq h_X(\vartheta_n \beta + \varepsilon) + 2\varepsilon$  and the hypothesis of Theorem 7 is fulfilled whenever  $h_X(\vartheta_n \beta + \varepsilon) < 2\alpha - \vartheta_n \beta - 2\varepsilon$ . Because  $h_X$  is non-negative, this condition implies that  $2\varepsilon < 2\alpha - \vartheta_n \beta$ . Because  $h_X$  is increasing, it also implies that  $h_X(t) < t - 3\varepsilon, \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$  and the hypothesis of Theorem 9 is also fulfilled. We deduce the following theorem:

**Theorem 10.** *Let  $\varepsilon, \alpha$  and  $\beta$  be three non-negative real numbers such that  $\alpha \leq \beta$  and  $2\varepsilon < 2\alpha - \vartheta_n \beta$ . Let  $P$  be a finite set of points whose Hausdorff distance to a compact subset  $X$  is  $\varepsilon$  or less. Then, any  $(\alpha, \beta)$ -quasi Rips complex of  $P$  is homotopy equivalent to  $X^\eta$  for all  $\eta \in [2\alpha - \vartheta_n \beta - 2\varepsilon, \vartheta_n \beta + \varepsilon]$  whenever  $\alpha$  is an inert value of  $P$  and  $h_X(\vartheta_n \beta + \varepsilon) < 2\alpha - \vartheta_n \beta - 2\varepsilon$ .*

## 5.2. Connections with the critical function

In this section, we show that the class of shapes with an upper bounded convexity defect function are equivalent to the class of shapes with a lower bounded critical function. To make this idea precise, we need to recall the definition of critical functions instrumental in expressing sampling conditions for a class of shapes larger than those with a positive reach in [17]. Even though the distance function to  $X$  is not differentiable, it is possible to define a *generalized gradient function*  $\nabla_X : \mathbb{R}^n \setminus X \rightarrow \mathbb{R}^n$  that coincides with the usual gradient at points where  $d(\cdot, X)$  is differentiable and that vanishes precisely at points that are critical [17]. Specifically,

$$\nabla_X(y) = \frac{y - \text{Center}(\Gamma_X(y))}{d(y, X)}.$$

The critical function  $\chi_X : \mathbb{R}_+^* \rightarrow \mathbb{R}_+$  is defined by

$$\chi_X(t) = \inf_{d(y, X)=t} \|\nabla_X(y)\|.$$

Clearly, the critical function vanishes at  $t$  if and only if  $t$  is a critical value of the distance function. Thus, by Lemma 2, we have the equivalences:  $\chi_X(t) = 0 \iff c_X(t) = t \iff h_X(t) = t$ . The next two lemmas strengthen this fact. Our first lemma provides a lower bound on  $\chi_X$  at  $t$ , assuming an upper bound on  $c_X$  at  $t$ .

**Lemma 11.** *For all compact set  $X \subset \mathbb{R}^n$ , all  $0 \leq \mu \leq 1$  and all  $t \geq 0$ , the following implication holds:*

$$c_X(t) < (1 - \mu)t \implies \chi_X(t) > \mu.$$

*Proof.* Consider  $y \in \mathbb{R}^n$  such that  $d(y, X) = t$  and let us prove that  $\|\nabla_X(y)\| > \mu$ . Let  $\sigma = \Gamma_X(y)$  be the set of points in  $X$  with minimum distance to  $y$ ; see Figure 8, left. Suppose the smallest ball enclosing  $\sigma$  has center  $z$  and radius  $s$ . Since  $s \leq t$ , we get  $c_X(s) \leq c_X(t) < (1 - \mu)t$  and thus  $t - \|y - z\| \leq d(z, X) \leq c_X(s) < (1 - \mu)t$ . It follows that  $\|\nabla_X(y)\| = \frac{\|z - y\|}{t} > \mu$ .  $\square$

Next lemma can be thought of as a converse of the previous lemma, since it provides an upper bound on  $h_X$  over the interval  $[0, R]$ , assuming a lower bound on the critical function  $\chi_X$  over the interval  $(0, R)$ . It extends a result in [7] which says intuitively that the convex hull of point set  $\sigma \subset X$  cannot be too far away from a shape  $X$ , assuming  $\sigma$  can be enclosed in a ball of small radius  $t$  and  $X$  has a positive reach.

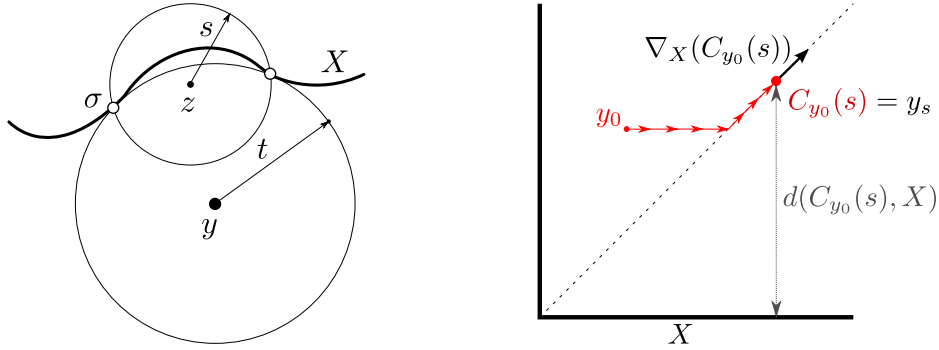


Figure 8: Notation for the proofs of Lemma 11 on the left and Lemma 12 on the right.

**Lemma 12.** Consider two real numbers  $\mu \in (0, 1]$  and  $R \geq 0$ . Let  $X \subset \mathbb{R}^n$  be a compact set such that  $\chi_X(t) \geq \mu$  for all  $t \in (0, R)$ . Then, for all  $0 \leq t \leq R$ , one has:

$$h_X(t) \leq \frac{1 + \mu(1 - \mu) - \sqrt{1 - \mu(2 - \mu) \left(\frac{t}{R}\right)^2}}{\mu(2 - \mu)} R.$$

*Proof.* Given  $\sigma \subset X$  with  $\text{Rad}(\sigma) \leq R$  and  $y_0 \in \text{Hull}(\sigma)$ , we establish an upper bound on  $d(y_0, X)$  expressed as a function of  $\text{Rad}(\sigma)$ .

The first step in the proof is to find a point  $y_T$  that is “sufficiently” far away from  $X$  by following an integral line of the generalized gradient  $\nabla_X$  that originates at  $y_0$ . Let  $X^c = \mathbb{R}^n \setminus X$ . The author in [32] established that there exists a continuous map  $\Phi_X : \mathbb{R}_+ \times X^c \rightarrow X^c$  such that:

$$\frac{d}{dt^+} \Phi_X(t, y) = \nabla_X(\Phi_X(t, y))$$

where  $\frac{d}{dt^+}$  denotes the right derivative. Hence,  $\Phi_X$  is a flow and the integral line  $t \mapsto \Phi_X(t, y_0)$  is a rectifiable curve starting at  $y_0$ . This curve either has infinite length or ends up at a critical point of the distance function to  $X$ . If for some  $T > 0$  the set  $\Phi_X([0, T], y_0)$  contains no critical point then it can be parameterized by a continuous function  $\mathcal{C}_{y_0} : [0, L] \rightarrow \mathbb{R}^n$  such that  $\mathcal{C}_{y_0}(0) = y_0$  and the length of  $\mathcal{C}_{y_0}([0, s])$  is  $s$ ; see Figure 8, right. Let us prove that under the assumption of the lemma we can choose  $T = R - d(y_0, X)$ . For all  $s < R - d(y_0, X)$ , we note that  $d(y_s, X) \leq d(y_0, X) + \|y_s - y_0\| \leq d(y_0, X) + s < R$  and therefore  $\chi_X(d(y_s, X)) \geq \mu$  which implies  $\|\nabla_X(y_s)\| \geq \mu$ . It follows that the integral line  $\mathcal{C}_{y_0}$  does not reach any critical point as

long as  $s < R - d(y_0, X)$  and  $\mathcal{C}_{y_0}$  can at least be parameterized on the interval  $[0, R - d(y_0, X)]$ . Hence, we can set  $T = R - d(y_0, X)$ . It has been established in [32, 18] that:

$$\forall s \in [0, L), \quad \frac{d}{ds^+} d(\mathcal{C}_{y_0}(s), X) = \|\nabla_X(\mathcal{C}_{y_0}(s))\|$$

Integrating over the interval  $[0, T]$ , we get

$$\frac{d(y_T, X) - d(y_0, X)}{T} \geq \mu.$$

Applying Lemma 1 with  $x = y_T$  and  $y = y_0$  gives  $d(y_T, \sigma)^2 \leq T^2 + \text{Rad}(\sigma)^2$  from which we deduce that

$$(d(y_0, X) + \mu T)^2 \leq d(y_T, X)^2 \leq d(y_T, \sigma)^2 \leq T^2 + \text{Rad}(\sigma)^2.$$

Plugging  $T = R - d(y_0, X)$ , setting  $\delta = \frac{d(y_0, X)}{R}$ ,  $\rho = \frac{\text{Rad}(\sigma)}{R}$  and rearranging this inequality gives us

$$\mu(2 - \mu)\delta^2 - 2(1 + \mu - \mu^2)\delta + 1 - \mu^2 + \rho^2 \geq 0.$$

Since  $\delta \leq 1$  we get  $\delta \leq \frac{1 + \mu(1 - \mu) - \sqrt{1 - \rho^2 \mu(2 - \mu)}}{\mu(2 - \mu)}$ , yielding the result.  $\square$

The upper bound on  $h_X$  is an arc of ellipse which tends to an arc of parabola as  $\mu \rightarrow 0$ ; see Figure 7, right. Note that since  $h_X(t) \leq t$  for all  $t$ , this upper bound is only relevant when under the diagonal. For  $\mu = 1$ , we get  $h_X(t) \leq R - \sqrt{R^2 - t^2}$  as in [7]. Equivalently, the graph of  $h_X$  is below the circle with radius  $R$  and center  $(0, R)$ .

### 5.3. Reconstructing shapes with a positive $\mu$ -reach

Shapes with a positive  $\mu$ -reach form a large class of objects, that unlike shapes with a positive reach, may possess sharp concave edges. Precisely, for  $0 < \mu \leq 1$ , authors in [17] define the  $\mu$ -reach of  $X$  as  $r_\mu(X) = \inf \{t > 0 \mid \chi_X(t) < \mu\}$ . The terminology comes from the fact that  $r_1(X)$  coincides with the usual reach of  $X$ .

Given a shape  $X$  whose  $\mu$ -reach is greater than or equal to  $R > 0$  and a finite point set  $P$  such that  $d_H(P, X) \leq \varepsilon$ , we compute the largest value of the ratio  $\frac{\varepsilon}{R}$  for which the Čech complex  $\mathcal{C}(P, \alpha)$  or the Rips complex  $\mathcal{R}(P, \alpha)$  provide a topologically correct reconstruction of  $X$  for a suitable value of the parameter  $\alpha$ . Computations were realized using a computer algebra system and details are skipped. In Appendix C, we give all the details when  $\mu = 1$ ,  $R = 1$  and  $n = +\infty$ .

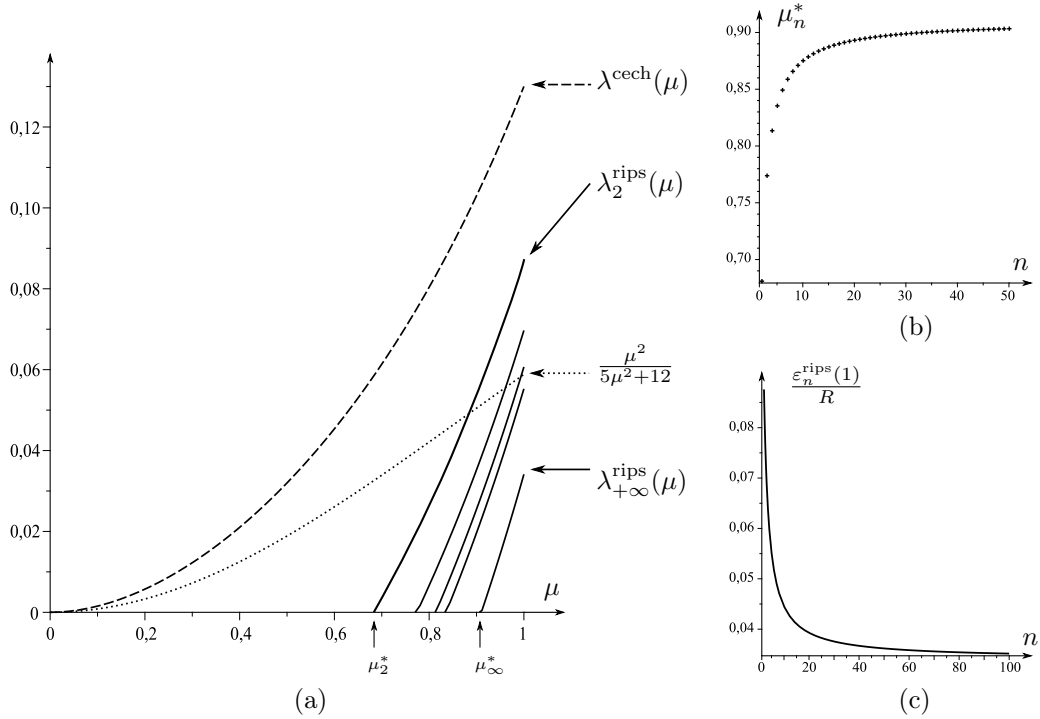


Figure 9: (a) Best ratios  $\frac{\varepsilon}{R}$  we can get for a correct reconstruction of a shape with a positive  $\mu$ -reach either with the Čech or Rips complex for  $n \in \{2, 3, 4, 5, +\infty\}$ ; comparison with the ratio  $\frac{\mu^2}{5\mu^2+12}$  obtained in [17]. (b) and (c)  $\mu_n^*$  and  $\lambda_n^{\text{rips}}(1)$  as functions of  $n$ .

*Reconstruction with the Čech complex.* Note that our assumption  $r_\mu(X) \geq R$  is equivalent to  $\chi_X(t) \geq \mu$  for all  $t \in (0, R)$ . It follows that a shape  $X$  with  $r_\mu(X) \geq R$  satisfies the hypothesis of Lemma 12 and therefore has a convexity defects function  $h_X$  upper bounded by a function that depends upon  $R$  and  $\mu$ . Plugging this upper bound in Theorem 9, we obtain that if  $\alpha + 3\varepsilon \leq R$  then the Čech complex  $\mathcal{C}(P, \alpha)$  is homotopy equivalent to  $X^\eta$  for all  $0 < \eta < R$  whenever the following inequality holds for all  $t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$ :

$$\frac{1 + \mu(1 - \mu) - \sqrt{1 - \mu(2 - \mu) \left(\frac{t}{R}\right)^2}}{\mu(2 - \mu)} R < t - 3\varepsilon.$$

Eliminating the square root, we can replace the above inequality by  $H_{\mu,\varepsilon}(t) < 0$  where  $H_{\mu,\varepsilon}(t)$  is a polynomial of degree 2 in  $t$ . It follows that the above condition holds whenever the absolute difference between the two roots  $t_\mu^1(\varepsilon) \leq$



$t_\mu^2(\varepsilon)$  of  $H_{\mu,\varepsilon}(t)$  is greater than  $2\varepsilon$ . When this happens, the admissible values of  $\alpha$  range in the interval  $I_\mu(\varepsilon) = [t_\mu^1(\varepsilon) - \varepsilon, t_\mu^2(\varepsilon) - 3\varepsilon]$ . The condition  $t_\mu^2(\varepsilon) - t_\mu^1(\varepsilon) > 2\varepsilon$  can be rewritten as the positivity of a polynomial of degree 2 in  $\varepsilon$  with two roots, one positive and one negative. Thus, the condition holds whenever  $\varepsilon$  is smaller than the positive root whose value divided by  $R$  is:

$$\lambda^{\text{cech}}(\mu) = \frac{-3\mu + 3\mu^2 - 3 + \sqrt{-8\mu^2 + 4\mu^3 + 18\mu + 2\mu^4 + 9 + \mu^6 - 4\mu^5}}{-7\mu^2 + 22\mu + \mu^4 - 4\mu^3 + 1}.$$

We thus get the following reconstruction theorem:

**Theorem 13.** *Consider a finite set of points  $P$  and a compact subset  $X$  whose  $\mu$ -reach  $R$  is positive. If*

$$d_H(P, X) \leq \varepsilon < \lambda^{\text{cech}}(\mu) R$$

*then  $\mathcal{C}(P, \alpha)$  is homotopy equivalent to  $X^\eta$  for all  $\eta \in (0, R)$  and all  $\alpha \in I_\mu(\varepsilon)$ .*

Interestingly,  $\lambda^{\text{cech}}(\mu)$  does not depend on the ambient dimension  $n$ . Plotting  $\lambda^{\text{cech}}(\mu)$  as a function of  $\mu$  (see Figure 9(a)), we observe that it is positive for all  $\mu \in (0, 1]$  and improves on the upper bound  $\frac{\mu^2}{5\mu^2+12}$  established in [17]. Still, for  $\mu = 1$ , we get  $\lambda^{\text{cech}}(1) = \frac{-3+\sqrt{22}}{13} \approx 0.13$  which is not as good as the value  $3 - \sqrt{8} \approx 0.17$  obtained in [34].

*Reconstruction with the Rips complex.* Combining Theorem 10 with  $\beta = \alpha$  and Lemma 12, we get that if  $\vartheta_n \alpha + \varepsilon \leq R$  then the Rips complex  $\mathcal{R}(P, \alpha)$  is homotopy equivalent to  $X^\eta$  for all  $0 < \eta < R$  whenever

$$\frac{1 + \mu(1 - \mu) - \sqrt{1 - \mu(2 - \mu) \left(\frac{\vartheta_n \alpha + \varepsilon}{R}\right)^2}}{\mu(2 - \mu)} R < 2\alpha - \vartheta_n \alpha - 2\varepsilon.$$

As before, we can eliminate the square root, replacing the above inequality by  $H_\mu(\varepsilon, \alpha) < 0$  where  $H_\mu(\varepsilon, \alpha)$  is a polynomial of degree 2 in  $\varepsilon$  and  $\alpha$ . Since we are looking for the greatest value of  $\varepsilon$  for which  $H_\mu(\varepsilon, \alpha) < 0$ , we may assume that  $\frac{\partial H_\mu(\varepsilon, \alpha)}{\partial \alpha} = 0$ . Plugging the value of  $\alpha$  for which  $\frac{\partial H_\mu(\varepsilon, \alpha)}{\partial \alpha} = 0$  in  $H_\mu(\varepsilon, \alpha)$ , we get a polynomial of degree 2 in  $\varepsilon$  whose greatest root  $\varepsilon_n^{\text{rips}}(\mu)$  gives the supremum of  $\varepsilon$  for which the above inequality holds. Setting  $\lambda_n^{\text{rips}}(\mu) = \frac{\varepsilon_n^{\text{rips}}(\mu)}{R}$  and letting  $\alpha_n^{\text{rips}}(\mu)$  be the value of  $\alpha$  for which  $H_\mu(\varepsilon_n^{\text{rips}}(\mu), \alpha) = 0$ , we get the following reconstruction theorem:

**Theorem 14.** *Consider a finite set of points  $P$  and a compact subset  $X$  whose  $\mu$ -reach  $R$  is positive. If*

$$d_H(P, X) \leq \varepsilon < \lambda_n^{\text{rips}}(\mu) R$$

*then  $\mathcal{R}(P, \alpha_n^{\text{rips}}(\mu))$  is homotopy equivalent to  $X^n$  for all  $n \in (0, R)$ .*

Using a computer algebra system, we obtain

$$\lambda_n^{\text{rips}}(\mu) = \frac{-\vartheta_n - \vartheta_n \mu + \vartheta_n \mu^2 + \sqrt{-4\mu^2 + 8\mu + 4\vartheta_n \mu^2 - 8\vartheta_n \mu - \vartheta_n^2 \mu^2 + 2\vartheta_n^2 \mu + \vartheta_n^2}}{\mu(2-\mu)(\vartheta_n+2)}$$

Plotting  $\lambda_n^{\text{rips}}(\mu)$  as a function of  $\mu$ , we observe that the ratio is only positive on a subinterval  $(\mu_n^*, 1]$  of  $(0, 1]$ ; see Figure 9(a). Hence, we can only guarantee that Rips complexes provide a correct reconstruction for shapes with a positive  $\mu$ -reach when  $\mu > \mu_n^*$ . In Figure 9(b), we plotted  $\mu_n^*$  as a function of  $n$ .  $\mu_n^*$  increases with  $n$  and we were able to prove using a computer algebra system that  $\mu_n^*$  tends to  $\sqrt{2\sqrt{2}-2} \approx 0.91$  as  $n \rightarrow +\infty$ . In Figure 9(c), we plotted  $\lambda_n^{\text{rips}}(1)$  as a function of  $n$ . It decreases with  $n$  and similarly, we proved that  $\lim_{n \rightarrow +\infty} \lambda_n^{\text{rips}}(1) = \frac{2\sqrt{2-\sqrt{2}}-\sqrt{2}}{2+\sqrt{2}} \approx 0.034$  and  $\lim_{n \rightarrow +\infty} \frac{\alpha_n^{\text{rips}}(1)}{R} = 1 - \frac{\sqrt{2}}{2} + \left(1 - \frac{3\sqrt{2}}{4}\right) \sqrt{2-\sqrt{2}} \approx 0.25$ .

## 6. Conclusion

Our work shows that Rips complexes can indeed provide topologically correct approximations of shapes.

An appealing aspect of Rips complexes is that they can be defined in metric spaces. Indeed, all we need is a notion of distances between points. Thus, it would be tempting to extend our results to metric spaces as well such as Riemannian manifolds,  $L^p$  spaces or abstract metric spaces. A glance at the proof of our main result (Theorem 7) reveals the centrality of Jung's Theorem that relates the diameter and radius of any subset of an Euclidean space. A first step towards such an extension would be to replace Jung's theorem by an axiom in the considered metric space. For instance, in  $L^\infty$  spaces the diameter of a set is always twice its radius. Hence, Čech and Rips complexes coincide and Theorem 7 degenerates into a trivial form. Besides replacing Jung's theorem by an axiom, we would also need to adapt our definitions of convexity defects functions. The definition of  $c_X$  is purely

metric. The definition of  $h_X$  requires, besides a metric, a notion of (local) convex hull which can be defined from geodesics for a large class of metric spaces (see the definition of *complete length spaces* in [28]).

Lemma 17 says that the hypotheses of Theorem 7 are stable under small metric perturbations. This seems to indicate that our relaxed definition of quasi Rips complexes (unlike the usual notion of Rips complexes) should allow us to apply Theorem 7 in the context of Gromov-Hausdorff distances.

Another natural extension would be to consider local measures of convexity defects and define sampling densities adapted to the local geometry of the sampled set in the spirit of [2].

## 7. Acknowledgment

We gratefully acknowledge Xavier Goaoc and Martin Tancer for fruitful discussions.

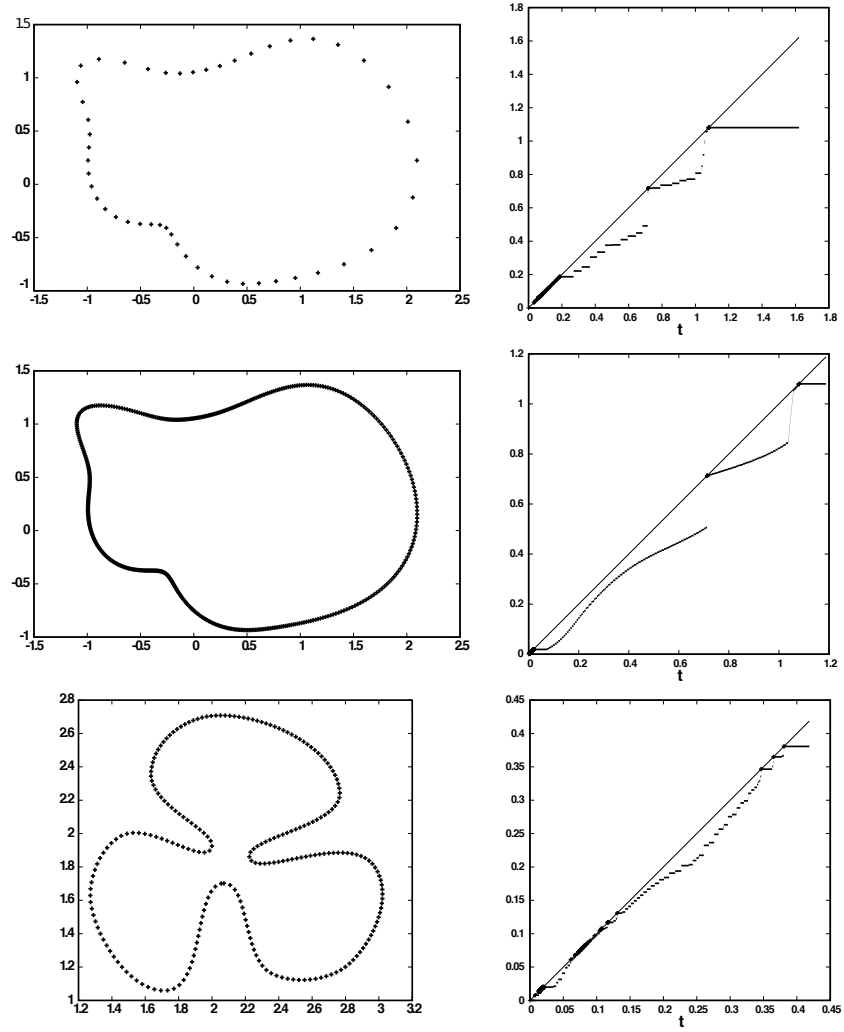


Figure 10: Various point sets  $P$  on the left and corresponding convexity defects function  $c_P$  on the right.

## Appendix A. Smallest enclosing balls

In this appendix, we establish the existence and uniqueness of the smallest ball enclosing a non-empty bounded set  $\sigma$  of  $\mathbb{R}^n$ . We then prove that the radius and center of this smallest enclosing ball are stable under perturbation of  $\sigma$ .

**Lemma 15.** *The smallest ball enclosing a non-empty bounded set of  $\mathbb{R}^n$  exists and is unique.*

*Proof.* Let  $\sigma$  be a non-empty bounded set of  $\mathbb{R}^n$ . We first establish the existence of a smallest ball enclosing  $\sigma$ . Given a point  $y \in \mathbb{R}^n$  and a real number  $s \geq 0$ , we first prove that the set  $\mathcal{B}(y, s)$  of closed balls passing through  $y$  and with radius  $s$  or less is compact. Indeed, representing a closed ball with center  $z$  and radius  $r$  by point  $(z, r)$  in  $\mathbb{R}^{n+1}$ , we can write

$$\mathcal{B}(y, s) = \{(z, r) \in \mathbb{R}^{n+1} \mid \|z - y\| \leq r \leq s\},$$

which is closed by definition and bounded since for all balls  $(z_0, r_0)$  and  $(z_1, r_1)$  in  $\mathcal{B}(y, s)$ , we have  $\|z_0 - z_1\| + |r_0 - r_1| \leq 3s$ . The set of closed balls containing  $\sigma$  and whose radii are smaller than or equal to the diameter of  $\sigma$  is

$$\mathcal{B}(\sigma) = \bigcap_{y \in \sigma} \mathcal{B}(y, \text{Diam}(\sigma)).$$

This set is non-empty and compact and therefore, the continuous map  $(z, r) \mapsto r$  on  $\mathcal{B}(\sigma)$  is bounded below and attains its infimum. The uniqueness is easy to establish by contradiction, as explained in [35].  $\square$

**Lemma 16.** *For every non-empty bounded subsets  $\sigma$  and  $\sigma'$  of  $\mathbb{R}^n$  such that  $d_H(\sigma, \sigma') \leq \varepsilon$ , we have  $|\text{Rad}(\sigma) - \text{Rad}(\sigma')| \leq \varepsilon$  and  $\|\text{Center}(\sigma) - \text{Center}(\sigma')\| \leq \sqrt{2\varepsilon \text{Rad}(\sigma) + \varepsilon^2}$ .*

*Proof.* Writing  $B$  for the smallest ball enclosing  $\sigma$ , we have  $\sigma' \subset \sigma^\varepsilon \subset B^\varepsilon$ , showing that  $\text{Rad}(\sigma') \leq \text{Rad}(\sigma) + \varepsilon$ . For the second part of the lemma, set  $z = \text{Center}(\sigma)$ ,  $z' = \text{Center}(\sigma')$ ,  $r = \text{Rad}(\sigma)$  and  $r' = \text{Rad}(\sigma')$ ; see Figure A.11. Suppose  $z \neq z'$  for otherwise the result is clear and consider the hyperplane  $H$  passing through  $z$  and orthogonal to the segment  $zz'$ . Let  $\xi = \bar{\sigma} \cap \partial B$ . By construction,  $\xi$  is closed and has the same smallest enclosing

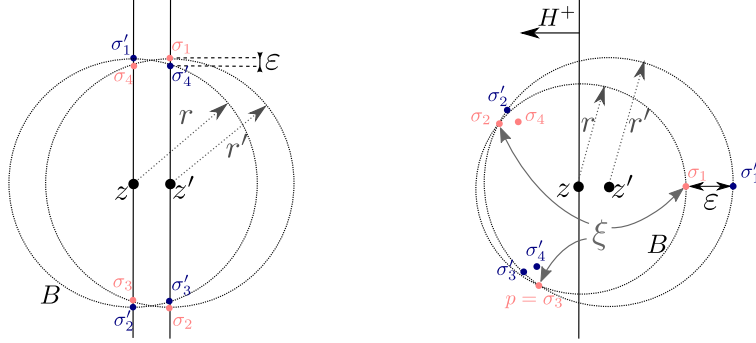


Figure A.11: Left: The bound on  $\|\text{Center}(\sigma) - \text{Center}(\sigma')\|$  in Lemma 16 is tight. Right: Notation for the proof of Lemma 16.

ball as  $\sigma$ . Thus,  $z \in \text{Hull}(\xi)$  and the closed half-space  $H^+$  that  $H$  bounds and which avoids  $z'$  intersects  $\xi$ . Let  $p$  be a point in this intersection. By choice of  $p$  in  $H^+$ , the triangle  $pzz'$  is obtuse at vertex  $z$  and  $\|z' - p\|^2 \geq \|z - p\|^2 + \|z' - z\|^2$ . Using  $\|z - p\| = r$  and  $\|z' - p\| \leq r' + \varepsilon$  we obtain  $\|z' - z\|^2 \leq (r' + \varepsilon)^2 - r^2$ . Interchanging the roles of  $\sigma$  and  $\sigma'$  yields:

$$\|z' - z\|^2 \leq \min\{(r' + \varepsilon)^2 - r^2, (r + \varepsilon)^2 - r'^2\}.$$

Note that  $(r' + \varepsilon)^2 - r^2 \leq (r + \varepsilon)^2 - r'^2$  if and only if  $r \geq r'$ . Considering in turn each of the two cases  $r \leq r'$  and  $r' \leq r$ , we get the desired inequality.  $\square$

## Appendix B. Hypotheses of Theorem 7 are stable

In this appendix, we establish the stability of hypotheses of Theorem 7 under small perturbations of the point set  $P$ . Given a point set  $P \subset \mathbb{R}^n$ , we say that a map  $f : P \rightarrow \mathbb{R}^n$  is an  $\varepsilon$ -small perturbation of  $P$  if  $f$  is injective and  $\|p - f(p)\| \leq \varepsilon$  for all points  $p \in P$ . Given a simplicial complex  $K$ , we define the simplicial complex  $f(K) = \{f(\sigma) \mid \sigma \in K\}$ .

**Lemma 17.** *Let  $P \subset \mathbb{R}^n$  be a finite set of points. Consider two real numbers  $\beta \geq \alpha \geq 0$  such that*

$$c_P(\vartheta_n \beta) < 2\alpha - \vartheta_n \beta$$

*and suppose moreover that  $\alpha$  is an inert value of  $P$ . Then, there exist  $\varepsilon > 0$  and  $\beta' > \beta$  such that for all  $\varepsilon$ -small perturbations  $f$  of  $P$ , we have:*

$$(i) \ c_{f(P)}(\vartheta_n \beta') < 2\alpha - \vartheta_n \beta';$$

- (ii)  $\mathcal{C}(f(P), \alpha) = f(\mathcal{C}(P, \alpha))$ ;
- (iii) if Flag  $G$  is an  $(\alpha, \beta)$ -quasi Rips complex of  $P$ , then Flag  $f(G)$  is an  $(\alpha, \beta')$ -quasi Rips complex of  $f(P)$ .

*Proof.* Let us establish (i). For this, set  $t = \vartheta_n \beta$  and define  $\bar{t} = \min\{\text{Rad}(\sigma) \mid \emptyset \neq \sigma \subset P \text{ and } \text{Rad}(\sigma) > t\}$ . By construction,  $\bar{t} > t$ . Lemma 4 ensures that for all subsets  $P' \subset \mathbb{R}^n$  within Hausdorff distance  $\varepsilon$  from  $P$  and for all  $t' \geq 0$ , the following implication holds:

$$c_{P'}(t') < c_P(t' + \varepsilon) + \sqrt{2t'\varepsilon + \varepsilon^2} + \varepsilon.$$

By assumption, we have  $2\alpha - t - c_P(t) > 0$ . By choosing  $\varepsilon > 0$  small enough, we can always find  $t' > t$  such that (1)  $t' + \varepsilon < \bar{t}$ , (2)  $2\alpha - t' - c_P(t) > 0$  and (3)  $\sqrt{2t'\varepsilon + \varepsilon^2} + \varepsilon \leq \frac{2\alpha - t' - c_P(t)}{2}$ . Since  $c_{P'}(t' + \varepsilon) = c_P(t)$ , it follows that

$$c_{P'}(t') < c_P(t) + \frac{2\alpha - t' - c_P(t)}{2} < 2\alpha - t'$$

and (i) is proved with  $\beta' = t'/\vartheta_n$ . By choosing  $\varepsilon > 0$  small enough, we can always assume that in addition to conditions (1), (2) and (3), we have (4)  $\varepsilon < \beta' - \beta$  and (5)  $\text{Rad}(\sigma) \notin [\alpha - \varepsilon, \alpha + \varepsilon]$  for all  $\emptyset \neq \sigma \subset P$ . Let  $f$  be an  $\varepsilon$ -small perturbation of  $P$ . Using Lemma 16 and condition (5), we get

$$\begin{aligned} \sigma \in \mathcal{C}(P, \alpha) &\Leftrightarrow \text{Rad}(\sigma) \leq \alpha \Leftrightarrow \text{Rad}(\sigma) \leq \alpha - \varepsilon \\ &\Rightarrow \text{Rad}(f(\sigma)) \leq \alpha \Leftrightarrow f(\sigma) \in \mathcal{C}(f(P), \alpha) \end{aligned}$$

and

$$\begin{aligned} f(\sigma) \in \mathcal{C}(f(P), \alpha) &\Leftrightarrow \text{Rad}(f(\sigma)) \leq \alpha \Rightarrow \text{Rad}(\sigma) \leq \alpha + \varepsilon \\ &\Leftrightarrow \text{Rad}(\sigma) \leq \alpha \Leftrightarrow \sigma \in \mathcal{C}(P, \alpha), \end{aligned}$$

yielding (ii). Consider a graph  $G$  whose flag complex is an  $(\alpha, \beta)$ -quasi complex and let  $p$  and  $q$  be two points of  $P$  such that  $\|f(p) - f(q)\| \leq 2\alpha$ . We have  $\|p - q\| \leq 2\alpha + 2\varepsilon$  and therefore using condition (5)  $\|p - q\| \leq 2\alpha$ . It follows that the edge  $\{p, q\}$  belongs to  $G$  and consequently the edge  $\{f(p), f(q)\}$  belongs to  $f(G)$ . Similarly, suppose  $\|f(p) - f(q)\| > 2\beta'$ . This implies that  $\|p - q\| > 2\beta' - 2\varepsilon > 2\beta$  by condition (4) and therefore the edge  $\{p, q\}$  does not belong to  $G$ . Hence, the edge  $\{f(p), f(q)\}$  does not belong to  $f(G)$ , showing (iii).  $\square$

## Appendix C. Reconstructing shapes with a positive reach

In this appendix, we redo computations of Section 5.3, setting  $\mu = 1$ ,  $R = 1$ ,  $n = +\infty$ . That is, we consider a shape  $X$  whose reach is greater than or equal to 1 and a finite point set  $P$  such that  $d_H(P, X) \leq \varepsilon$ .

*Reconstruction with the Čech complex.* Combining Theorem 9 and Lemma 12, we get that the Čech complex  $\mathcal{C}(P, \alpha)$  is homotopy equivalent to  $X^\eta$  for  $0 < \eta < 1$  whenever

$$1 - \sqrt{1 - t^2} < t - 3\varepsilon, \quad \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$$

which can be rewritten as

$$2t^2 - 2t(1 + 3\varepsilon) + 9\varepsilon^2 + 6\varepsilon < 0, \quad \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon].$$

This condition holds whenever the absolute difference between the two roots  $t^1(\varepsilon) \leq t^2(\varepsilon)$  of the polynomial in  $t$  is greater than  $2\varepsilon$ , that is, whenever  $0 > 13\varepsilon^2 + 6\varepsilon - 1$ . The supremum of  $\varepsilon$  for which the previous equation holds is  $\lambda^{\text{cech}}(1) = \frac{-3 + \sqrt{22}}{13} \approx 0.13$ . For  $\varepsilon < \lambda^{\text{cech}}(1)$ , the admissible values of  $\alpha$  range in the interval  $I_1(\varepsilon) = [t^1(\varepsilon) - \varepsilon, t^2(\varepsilon) - 3\varepsilon]$ . Plugging the expression of the roots into the two endpoints of the interval, we get

$$I_1(\varepsilon) = \left[ \frac{1}{2} + \varepsilon - \frac{\sqrt{1 - 6\varepsilon - 9\varepsilon^2}}{2}, \frac{1}{2} - \frac{3\varepsilon}{2} + \frac{\sqrt{1 - 6\varepsilon - 9\varepsilon^2}}{2} \right].$$

Note that when epsilon tends to  $\lambda^{\text{cech}}(1)$ , the interval  $I_1(\varepsilon)$  tends to the singleton  $\frac{8}{13} - \frac{\sqrt{22}}{26} \approx 0.44$ .

*Reconstruction with the Rips complex.* Combining Theorem 10 and Lemma 12, we get that the Rips complex  $\mathcal{R}(P, \alpha)$  is homotopy equivalent to  $X^\eta$  for all  $0 < \eta < R$  whenever

$$1 - \sqrt{1 - (\sqrt{2}\alpha + \varepsilon)^2} < 2\alpha - \sqrt{2}\alpha - 2\varepsilon.$$

which we can rewrote as

$$5\varepsilon^2 + 4(2 - \sqrt{2})\alpha^2 - 2(4 - 3\sqrt{2})\alpha\varepsilon + 4\varepsilon - 2(2 - \sqrt{2})\alpha < 0$$

Since we are looking for the greatest value of  $\varepsilon$  for which the above equation holds, we may assume that the partial derivative of the left side with respect



to  $\alpha$  vanishes, which gives  $4(2 - \sqrt{2})\alpha - (4 - 3\sqrt{2})\varepsilon - (2 - \sqrt{2}) = 0$ . Plugging  $\alpha = ((1 - \sqrt{2})\varepsilon + 1)/4$  in the above equation, we get

$$(10 + 7\sqrt{2})\varepsilon^2 + (8 + 6\sqrt{2})\varepsilon + \sqrt{2} - 2 < 0$$

The left side is a polynomial of degree 2 in  $\varepsilon$  whose greatest root  $\varepsilon_{+\infty}^{\text{rips}}(1) = \frac{2\sqrt{2-\sqrt{2}}-\sqrt{2}}{2+\sqrt{2}} \approx 0.034$  gives the supremum of  $\varepsilon$  for which the above inequality holds. The corresponding value of  $\alpha$  is

$$\alpha_{+\infty}^{\text{rips}}(1) = 1 - \frac{\sqrt{2}}{2} + \left(1 - \frac{3\sqrt{2}}{4}\right) \sqrt{2 - \sqrt{2}} \approx 0.25.$$

## References

- [1] N. Amenta, D. Attali, and O. Devillers. Complexity of Delaunay triangulation for points on lower-dimensional polyhedra. In *Proc. ACM-SIAM Sympos. Discrete Algorithms (SODA '07)*, New-Orleans, Louisiana, USA, January 7–9 2007.
- [2] N. Amenta and M. Bern. Surface reconstruction by Voronoi filtering. *Discrete and Computational Geometry*, 22(4):481–504, 1999.
- [3] N. Amenta, M. Bern, and D. Eppstein. The crust and the  $\beta$ -skeleton: Combinatorial curve reconstruction. *Graphical Models and Image Processing*, 60(2):125–135, 1998.
- [4] D. Attali.  $r$ -regular shape reconstruction from unorganized points. *Computational Geometry: Theory and Applications*, 10:239–247, 1998.
- [5] D. Attali and J.-D. Boissonnat. A linear bound on the complexity of the Delaunay triangulation of points on polyhedral surfaces. *Discrete and Computational Geometry*, 31(3):369–384, February 2004.
- [6] D. Attali, J.-D. Boissonnat, and A. Lieutier. Complexity of the Delaunay triangulation of points on surfaces: the smooth case. In *19th ACM Symposium on Computational Geometry*, pages 201–210, San-Diego, California, USA, June 2003.
- [7] D. Attali and A. Lieutier. Reconstructing shapes with guarantees by unions of convex sets. In *26th Ann. Sympos. Comput. Geom.*, pages 344–353, Snowbird, Utah, June 13-16 2010.
- [8] D. Attali, A. Lieutier, and D. Salinas. Efficient Data Structure for Representing and Simplifying Simplicial Complexes in High Dimension. In *Proc. 27th Ann. Sympos. Comput. Geom.*, pages 501–509, Paris, France, June 13-15 2011.
- [9] J. A. Barmak and E. G. Minian. Strong homotopy types, nerves and collapses. Technical report, arXiv:0907.2954v1, 2009.
- [10] A. Bjorner. Topological methods. In *Handbook of combinatorics (vol. 2)*, page 1850. MIT Press, 1996.

- [11] J. Boissonnat and F. Cazals. Smooth surface reconstruction via natural neighbour interpolation of distance functions. *Computational Geometry: Theory and Applications*, 22(1-3):185–203, 2002.
- [12] J. Boissonnat, O. Devillers, and S. Hornus. Incremental construction of the Delaunay triangulation and the Delaunay graph in medium dimension. In *Proceedings of the 25th annual symposium on Computational geometry*, pages 208–216. ACM, 2009.
- [13] J. Boissonnat and A. Ghosh. Manifold reconstruction using tangential Delaunay complexes. In *Proceedings of the 2010 annual symposium on Computational geometry*, pages 324–333. ACM, 2010.
- [14] G. Carlsson and V. de Silva. Topological approximation by small simplicial complexes. *preprint*, 2003.
- [15] E. Chambers, V. de Silva, J. Erickson, and R. Ghrist. Vietoris–rips complexes of planar point sets. *Discrete and Computational Geometry*, pages 1–16, 2010.
- [16] E. Chambers, V. De Silva, J. Erickson, and R. Ghrist. Vietoris–rips complexes of planar point sets. *Discrete & Computational Geometry*, 44(1):75–90, 2010.
- [17] F. Chazal, D. Cohen-Steiner, and A. Lieutier. A sampling theory for compact sets in Euclidean space. *Discrete and Computational Geometry*, 41(3):461–479, 2009.
- [18] F. Chazal and A. Lieutier. The  $\lambda$ -medial axis. *Graphical Models*, 67(4):304–331, 2005.
- [19] F. Chazal and A. Lieutier. Smooth Manifold Reconstruction from Noisy and Non Uniform Approximation with Guarantees. *Computational Geometry: Theory and Applications*, 40:156–170, 2008.
- [20] F. Chazal and S. Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proc. ACM Symposium on Computational Geometry*, pages 232–241, 2008.
- [21] V. De Silva. A weak characterisation of the Delaunay triangulation. *Geometriae Dedicata*, 135(1):39–64, 2008.

- [22] V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7:339–358, 2007.
- [23] T. Dey. *Curve and surface reconstruction: algorithms with mathematical analysis*. Cambridge Univ Pr, 2007.
- [24] H. Edelsbrunner. The union of balls and its dual shape. In *Proceedings of the ninth annual symposium on Computational geometry*, pages 218–231. ACM, 1993.
- [25] H. Edelsbrunner and J. Harer. *Computational topology: an introduction*. Amer Mathematical Society, Providence, Rhode Island, 2010.
- [26] H. Edelsbrunner and E. Mücke. Three-dimensional alpha shapes. In *Proceedings of the 1992 workshop on Volume visualization*, pages 75–82. ACM, 1992.
- [27] H. Edelsbrunner and N. Shah. Triangulating topological spaces. In *International Journal of Computational Geometry and Applications*, volume 7, pages 365–378. Citeseer, 1997.
- [28] M. Gromov, M. Katz, P. Pansu, and S. Semmes. *Metric structures for Riemannian and non-Riemannian spaces*. Birkhauser, 2006.
- [29] K. Grove. Critical point theory for distance functions. In *Proc. of Symposia in Pure Mathematics*, volume 54, pages 357–386, 1993.
- [30] J. Hausmann. On the Vietoris-Rips Complexes and a Cohomology Theory for Metric Spaces. In *Prospects in topology: proceedings of a conference in honor of William Browder*, page 175. Princeton Univ Pr, 1995.
- [31] J. Latschev. Vietoris-Rips complexes of metric spaces near a closed Riemannian manifold. *Archiv der Mathematik*, 77(6):522–528, 2001.
- [32] A. Lieutier. Any open bounded subset of  $\mathbb{R}^n$  has the same homotopy type as its medial axis. *Computer-Aided Design*, 36(11):1029–1046, 2004.
- [33] J. Munkres. *Elements of algebraic topology*. Perseus Books, 1993.
- [34] P. Niyogi, S. Smale, and S. Weinberger. Finding the Homology of Submanifolds with High Confidence from Random Samples. *Discrete Computational Geometry*, 39(1-3):419–441, 2008.

- [35] E. Welzl. Smallest enclosing disks (balls and ellipsoids). In H. Maurer, editor, *New Results and New Trends in Computer Science*, volume 555, pages 359–370. Springer, 1991.
- [36] A. Zomorodian. Fast construction of the Vietoris-Rips complex. *Computers & Graphics*, 2010.
- [37] A. Zomorodian. The tidy set: a minimal simplicial set for computing homology of clique complexes. In *Proceedings of ACM symposium of computational geometry*, 2010.