Vietoris-Rips complexes also provide topologically correct reconstructions of sampled shapes*

Dominique Attali[†] André Lieutier[‡] David Salinas[†]

Submitted to SoCG 2011 — November 30, 2010

Abstract

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 improve constants when X has a positive μ-reach. Most importantly, our work shows that Rips complexes can also be used to provide topologically correct reconstruction of shapes. This may be of some computational interest in high dimension.

*This work is partially supported by ANR Project *GIGA* ANR-09-BLAN-0331-01.

[†]Gipsa-lab – CNRS UMR 5216, Grenoble, France. `Firstname.Lastname@gipsa-lab.grenoble-inp.fr`
[‡]Dassault systèmes, Aix-en-Provence, France. `andre.lieutier@3ds.com`

1 Introduction

In this paper, we formulate conditions under which Rips complexes reconstruct shapes using measures of how far the shape is from being convex.

Motivation. The problem of reconstructing shapes from point clouds arises in many fields, including computer graphics and machine learning [ABE98, Dey07]. Maybe one of the simplest reconstruction method is to output an α-offset of the sample points, that is, the union of balls centered at the sample with radius α. Assuming the shape is a smooth manifold [NSW08, CL08] or more generally has a positive μ-reach [CCSL09], 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 α. Topologically, this is equivalent to computing the α-shape [Ede93, EM92] of the sample points, which can be obtained by first building the Delaunay triangulation and then keeping simplices that fit in a ball of radius α.

This approach works well for point clouds in three-dimensional space which have Delaunay triangulations of affordable size [AB04, ABL03]. But, as the dimension of the ambient space increases, the size of the Delaunay triangulation explodes [AAD07] and other strategies must be found. If the data points lie on a low-dimensional submanifold, it seems reasonable to ask that the building of the reconstruction depends only upon the intrinsic dimension of the data. This motivated de Silva [DS08] to introduce *Witness complexes* and Boissonnat and Gosh [BG10] to define *tangential Delaunay complexes*. For medium dimensions, Boissonnat and al. [BDH09] 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 α, the *Vietoris-Rips complex* is the simplicial complex whose simplices are subsets of points in P with diameter at most 2α. Rips complexes are examples of *flag complexes* and as so, 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 [Zom10a, Zom10b] and edge contractions [ALS]. 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 α consists of all simplices spanned by points in P that fit in a ball of radius α. The construction is similar to that of α-shapes, but without the restriction that simplices belong to the Delaunay triangulation. The Čech complex at scale α is homotopy equivalent to α-offset and therefore also possesses the ability to reproduce the topology of the sampled shape. This property was used by Chazal and Oudot [CO08] 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 the Rips complexe captures the topology of the shape. In a different setting, it has been proved in [Hau95, Lat01] 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.

Partially related to our work, we should mention [CdSEG10] 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 [Att98, AB99, BC02, NSW08, CL08]. 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 [CCSL09] introduce a new characterization of the feature size, the μ-*reach*, which allows 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 curvatures measure how far an object is from being locally flat. In Section 4, 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 μ-reach. Most importantly, this new framework allows us to prove that Rips complexes also provide topologically correct reconstruction, assuming shapes have a positive μ-reach, for μ sufficiently large. For this, we first find conditions under which Rips complexes collapse to Čech complexes in Section 3.

By lack of space, we put the proofs of some of the lemmas in Appendix C.

2 Preliminaries

In this section, we introduce the definitions and tools we need to state and prove our results.

2.1 Metric space and distances

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 $d_H(Y \mid X) = \sup_{y \in Y} d(y, X)$ 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 $d_H(Y \mid X) \leq \varepsilon$ if and only if Y is contained in the ε-offset $X^\varepsilon = \{y \in \mathbb{R}^n \mid d(y, X) \leq \varepsilon\}$. The *Hausdorff distance* between X and Y is $d_H(X, Y) = \max\{d_H(X \mid Y), d_H(Y \mid X)\}$. 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.

2.2 Smallest enclosing ball

We begin by stating that the smallest ball enclosing a non-empty bounded set is well-defined and then list some useful properties of its center and radius. Recall that the *diameter* of a subset σ of $\mathbb{R}^n$ is the supremum of distances between pairs of points in σ, which we denote as $\mathrm{Diam}(\sigma) = \sup_{p,q \in \sigma} \|p - q\|$. A subset σ is said to be *bounded* if its diameter is finite.

Lemma 1. *The smallest ball enclosing a non-empty bounded set of $\mathbb{R}^n$ exists and is unique.*

Given a non-empty bounded subset σ of $\mathbb{R}^n$, we denote the center and the radius of the smallest ball enclosing σ respectively by $\mathrm{Center}(\sigma)$ and $\mathrm{Rad}(\sigma)$. Writing $\mathrm{Hull}(X)$ for the convex hull of X and $\mathrm{Cl}\, X$ for the closure of X, it is not hard to check (by contradiction) that $\mathrm{Center}(\sigma) \in \mathrm{Cl}\,\mathrm{Hull}(\sigma)$.

Lemma 2. *For any non-empty bounded subset $\sigma \subset \mathbb{R}^n$, any point $x \in \mathbb{R}^n$ and any point $y \in \mathrm{Hull}(\sigma)$, we have that $d(x,\sigma)^2 \leq \|x-y\|^2 + \mathrm{Rad}(\sigma)^2 - \|y - \mathrm{Center}(\sigma)\|^2$.*

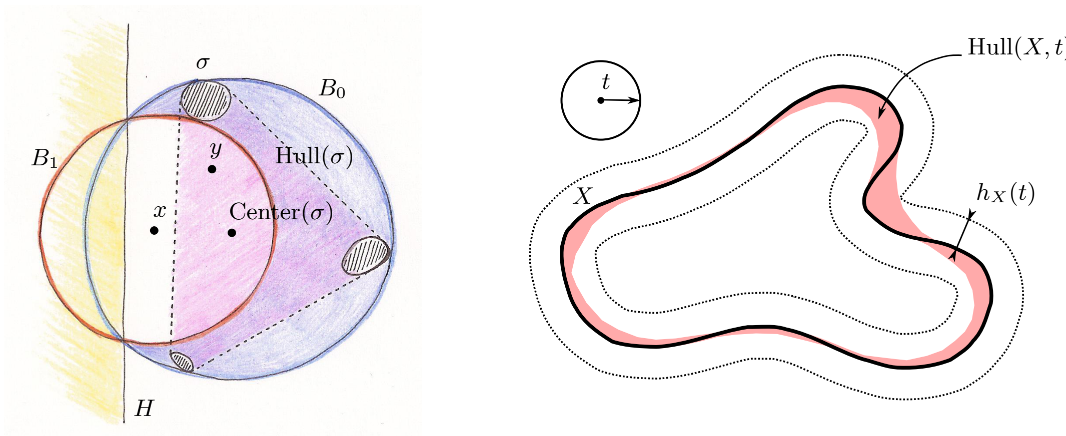

Figure 1: Left: Notations for the proof of Lemma 2. Right: Smallest offset of X containing $\mathrm{Hull}(X,t)$.

Proof. Suppose $d(x,\sigma) > \|x-y\|$ for otherwise the result is clear. Let B_0 be the smallest ball enclosing σ and let B_1 be the largest ball centered at x and whose interior does not intersect σ; see Figure 1, left. 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 be the set of points whose power distance to B_0 is at most as large as the power distance to B_1. H is a closed half-space which contains the set difference $B_0 \setminus B_1$. In particular, it contains σ and any point $y \in \mathrm{Hull}(\sigma)$. Thus, $\pi_{B_0}(y) \leq \pi_{B_1}(y)$ and the result follows. □

2.3 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 τ a *face* of σ and σ a *coface* of τ. 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, $\mathrm{Vert}(K) = \bigcup_{\sigma \in K} \sigma$. 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, $\mathrm{Shd}\, K = \bigcup_{\sigma \in K} \mathrm{Hull}(\sigma)$, not to be confused with $|K|$, the underlying space of a geometric realization of K; see [Mun93]. If N is the cardinal of the vertex set $\mathrm{Vert}(K)$ of K and if $f : \mathrm{Vert}(K) \to \mathbb{R}^N$ sends $\mathrm{Vert}(K)$ to an affinely independant set $f(\mathrm{Vert}(K))$, then $|K| = \bigcup_{\sigma \in K} \mathrm{Hull}(f(\sigma))$ (up to a homeomorphism). Generally, $|K|$ and $\mathrm{Shd}\, K$ are not homeomorphic since the relative interiors of the convex hulls of two different simplices of K may overlap.

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 may change from one author to another.

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, $\mathcal{C}(P,t) = \{\sigma \mid \emptyset \neq \sigma \subset P, \mathrm{Rad}(\sigma) \leq t\}$. Equivalently, a k-simplex $\{p_0, \ldots, p_k\}$ 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 $\mathrm{Nrv}\, F = \{G \subset F \mid \bigcap G \neq \emptyset\}$ 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 [Bjo96, ES97] implies that the Čech complex $\mathcal{C}(P, t)$ is homotopy equivalent to the union of these balls, that is, $|\mathcal{C}(P, t)| \simeq P^t$.

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 $2t$, $\mathcal{R}(P, t) = \{\sigma \mid \emptyset \neq \sigma \subset P, \mathrm{Diam}(\sigma) \leq 2t\}$. For simplicity, we refer to $\mathcal{R}(P, t)$ as the Rips complex. Recall that the *flag complex* of a graph G, denoted $\mathrm{Flag}\, G$, 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, $\mathcal{R}(P, t) = \mathrm{Flag}\left(\mathcal{C}(P, t)^{(1)}\right)$. Generally, $\mathcal{R}(P, t)$ and $\mathcal{C}(P, t)$ do not share the same topology. 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 $|\mathcal{R}(P, t)| \simeq P^t$. Along the way, we will need a result in [dSG07] which is a consequence of Jung's Theorem and which says that there is chain of inclusion

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

3 Condition under which Rips complexes and variants deformation retract to Čech complexes

In this section, we introduce 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 these functions, we will be able in Section 3.3 to formulate a condition which suffices to guarantee that Rips complexes of a finite set of points P deformation retracts to Čech complexes of P using a new kind of collapses described in Section 3.2. The condition refers only to the point set P.

3.1 Convexity defects measures

To avoid lengthy sentences, we adopt the convention that X 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. Recalling that $\mathrm{Hull}(X)$ denotes the convex hull of X, we first extend the notion of convex hull. We define the convex hull of X at scale t as the subset (see Figure 1, right)

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

Note that if P is a finite set of points, then $\mathrm{Hull}(P, t)$ is the shadow of the Čech complex $\mathcal{C}(P, t)$. Similarly, we define the set of centers of X at scale t as the subset:

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

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

We start with a remark. Because X is a subset of both $\mathrm{Hull}(X, t)$ and $\mathrm{Centers}(X, t)$, the two one-sided Hausdorff distances $d_H(X \mid \mathrm{Hull}(X, t))$ and $d_H(X \mid \mathrm{Centers}(X, t)$ vanish. It follows that we could have used in the above definition two-sided Hausdorff distances instead of one-sided Hausdorff distances. We now list a few basic properties of functions h_X and c_X (see Figure 4). First, h_X and c_X both vanish at 0, are increasing in the interval $[0, \mathrm{Rad}(X)]$ and become constant above $\mathrm{Rad}(X)$. Since $\mathrm{Center}(\sigma) \in \mathrm{Cl}\,\mathrm{Hull}(\sigma)$, 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 α, the following three conditions are equivalent: (1) $h_X(t) \leq \alpha$; (2) $\mathrm{Hull}(X, t) \subset X^\alpha$; (3) $[\mathrm{Rad}(\sigma) \leq t \implies \mathrm{Hull}(\sigma) \subset X^\alpha]$ for all $\sigma \subset X$. In particular, we get that $h_X(t) \leq t$ for all $t \geq 0$ since $\mathrm{Rad}(\sigma) \leq t \implies \mathrm{Hull}(\sigma) \subset \sigma^t$ by Lemma 2 applied for $x = y$.

Intuitively, h_X and c_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 $X \subset \mathbb{R}^n$ is convex if and only if $h_X = 0$. If X is compact, then X is convex if and only if $c_X = 0$. The two convexity functions h_X and c_X will play a different role. While c_P is all we need to study the Rips complex of a finite point P in Section 3.3, it turns out that h_X is more stable than c_X and will be used in Section 4.2 to express sampling conditions in reconstruction theorems.

3.2 Collapses

This section describes collapses that will be useful to deformation retraction Rips complexes to Čech complexes in the next section.

First, we need some definitions. Let σ be a simplex of the simplicial complex K. The *star* of σ in K, denoted $\mathrm{St}_K(\sigma)$, is the collection of simplices of K having σ as a face. The closure of $\mathrm{St}_K(\sigma)$ is denoted $\overline{\mathrm{St}}_K(\sigma)$; it is the smallest simplicial complex containing $\mathrm{St}_K(\sigma)$. The *link* of σ in K, denoted $\mathrm{Lk}_K(\sigma)$, is the collection of simplices of K lying in $\overline{\mathrm{St}}_K(\sigma)$ that are disjoint from σ. Given two simplicial complexes K and L, the smallest simplicial complex containing all the simplices of the form $\kappa \cup \lambda$ where $\kappa \in K$ and $\lambda \in L$ is called the *join* of K and L and is denoted by $K * L$. A simplicial complex K is said to be a *cone* if it contains a vertex o such that the following implication holds: $\sigma \in K \implies \sigma \cup \{o\} \in K$. Equivalently, a *cone* is the join $o * L$ of a (possibly empty) simplicial complex L and a vertex $o \notin L$. 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.

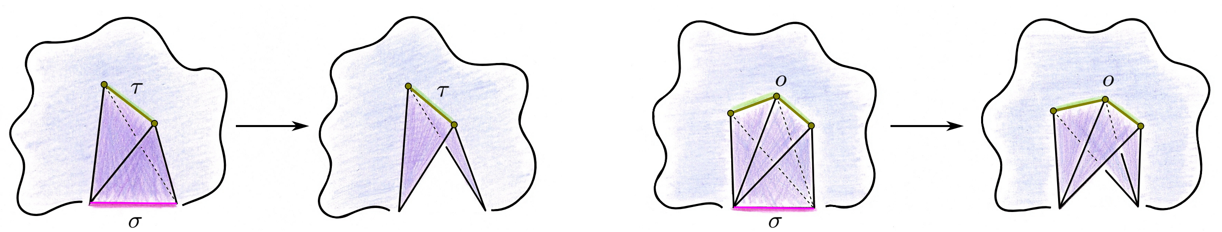

Figure 2: Left: In a classical collapse, the link of σ has a unique inclusion-maximal simplex τ. Right: in an extended collapse, the link of σ is a cone with apex o.

Given a simplicial complex K, we are interested in the operation that removes the entire star of a simplex $\sigma \in K$ (see Figure 2). Provided that there is a unique inclusion-maximal simplex $\tau \neq \sigma$ in the star of σ, it is well-known that $|K|$ deformation retracts to $|K \setminus \mathrm{St}_K(\sigma)|$ and the operation that removes $\mathrm{St}_K(\sigma)$ is then called a *collapse* [DEG99]. Following and extending what was done in [BM09], we still call a *collapse* the operation that removes $\mathrm{St}_K(\sigma)$ assuming the weaker condition that the link of σ is a cone. Our terminology finds its justification in the following lemma.

Lemma 3. *Let K be a simplicial complex and let σ be a simplex of K. If the link of σ is a cone, then $|K|$ deformation retracts to $|K \setminus \mathrm{St}_K(\sigma)|$.*

3.3 Almost Rips 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. As a consequence, we also state conditions under which a Čech complex deformation retracts to another one. Let us first define a 2-parameter family that contains prior Rips complexes as a subfamily:

Definition 2. *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 $\mathcal{R}(P, \alpha) \subset \operatorname{Flag} G \subset \mathcal{R}(P, \beta)$ an (α, β)-*almost Rips complex *of P.*

In other words, the simplicial complex $\operatorname{Flag} G$ is an (α, β)-almost Rips complex of P if and only if every pairs of points in P within distance 2α are connected by an edge in G and no edge of G has length larger than 2β. 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 (α, α)-almost Rips complex of P if and only if $K = \mathcal{R}(P, \alpha)$. To state our main theorem, it is convenient to define α to be an *inert value* of P if $\operatorname{Rad}(\sigma) \neq \alpha$ for all non-empty subsets $\sigma \subset P$. The finiteness of P implies that P has only finitely many non-inert values. Thus, assuming α to be inert is not a too restrictive hypothesis.

Theorem 1. *Let $P \subset \mathbb{R}^n$ be a finite set of points. For any real numbers $\beta \geq \alpha \geq 0$ such that α is an inert value of P and $c_P(\vartheta_{\mathrm{n}}\, \beta) < 2\alpha - \vartheta_{\mathrm{n}}\, \beta$, there exists a sequence of collapses from any (α, β)-almost Rips complex of P to the Čech complex $\mathcal{C}(P, \alpha)$.*

Proof. For $t \geq 0$, consider the simplicial complex $\mathcal{F}(t) = \mathcal{C}(P, t) \cap \operatorname{Flag} G$. By choice of α and β, there is chain of inclusions:

$$\mathcal{C}(P, \alpha) \ \subset \ \mathcal{R}(P, \alpha) \ \subset \ \operatorname{Flag} G \ \subset \ \mathcal{R}(P, \beta) \ \subset \ \mathcal{C}(P, \vartheta_{\mathrm{n}}\, \beta)$$

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

$$\mathcal{C}(P, \alpha) = \mathcal{C}_0 \ \subset \ \mathcal{C}_1 \ \subset \ \cdots \ \subset \ \mathcal{C}_k = \mathcal{C}(P, \vartheta_{\mathrm{n}}\, \beta).$$

For $0 < i < k$, let t_i be the smallest value of t such that $\mathcal{C}_i = \mathcal{C}(P, t)$ and set $\mathcal{F}_i = \mathcal{F}(t_i)$. Correspondingly, we get a 1-parameter family of simplicial complexes:

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

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

- $(\star)$ For all simplices $\sigma, \tau \subset P$, if $\operatorname{Rad}(\sigma) = \operatorname{Rad}(\tau)$ then $\operatorname{Center}(\sigma) = \operatorname{Center}(\tau)$;
- $(\star\star)$ 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 by showing that $\mathcal{C}_i$ collapses to $\mathcal{C}_{i-1}$ and $\mathcal{F}_i$ is either equal or collapses to $\mathcal{F}_{i-1}$ for all $0 < i \leq k$. Because of condition $(\star)$, all simplices in the difference $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$ share the same smallest enclosing ball $B(z_i, t_i)$. 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 σ_i. Note that the cofaces of σ_i in $\mathcal{C}_i$ have their vertices in the ball $B(z_i, t_i)$ and are thus faces of the simplex $\tau_i = \{p \in P \mid \|z_i - p\| \leq t_i\}$. In other words, the star of σ_i has a unique inclusion-maximal simplex τ_i. Let us prove that $\sigma_i \neq \tau_i$. The vertices of σ_i all lie on the sphere with center z_i. On the other hand, τ_i possesses at least a vertex, say o, at distance $c_P(t_i)$ or less from z_i. Since $c_P(t_i) \leq c_P(\vartheta_{\mathrm{n}}\, \beta) < \alpha \leq t_i$, the vertex o belongs to the interior of $B(z_i, t_i)$, $o \notin \sigma_i$, $\sigma_i \neq \tau_i$ and $\mathcal{C}_i$ collapses to $\mathcal{C}_{i-1}$.

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 σ_i in $\mathcal{F}_i$ is equal to the star of σ_i in $\mathcal{C}_i$ intersected with the flag of G and $\mathcal{F}_{i-1} = \mathcal{F}_i \setminus \mathrm{St}_{\mathcal{F}_i}(\sigma_i)$. Let us prove that the link of σ_i in $\mathcal{F}_i$ is a cone with apex o, which guarantees that $\mathcal{F}_i$ collapses to $\mathcal{F}_{i-1}$. Suppose η is a coface of σ_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 $p \in \eta \cup \{o\}$, $\|z_i - p\| \leq t_i$. 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 $\eta \cup \{o\} \in \mathrm{Flag}\, G$. Hence, $\eta \cup \{o\}$ belongs to $\mathcal{F}_i$. Since $o \notin \sigma_i$, it follows that the link of σ_i in $\mathcal{F}_i$ is a cone, which concludes the proof of Theorem 1 assuming generic conditions ($\star$) and ($\star\star$) instead of the condition $\mathrm{Rad}(\sigma) \neq \alpha$ for all non-empty subsets $\sigma \subset P$.

If P does not satisfy the generic conditions ($\star$) and ($\star\star$), we use Lemma 9 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 9 for some $\beta' > \beta$. Applying Theorem 1 to $f(P)$ with the values α and β', we get that there exists a sequence of collapses from the (α, β')-almost Rips complex $\mathrm{Flag}\, f(G) = f(\mathrm{Flag}(G))$ to the Čech complex $\mathcal{C}(f(P), \alpha) = f(\mathcal{C}(P, \alpha))$. Hence, the theorem also holds in the non-generic case. □

Choosing $\beta = \alpha$ in the theorem gives conditions under which $|\mathcal{R}(P, \alpha)| \simeq |\mathcal{C}(P, \alpha)| \simeq P^\alpha$. Figure 5, left provides a graphical representation of the hypothesis of the theorem. Slightly adapting the first part of the proof we get the following result:

Theorem 2. *Let $\beta \geq \alpha \geq 0$ and let P be a finite set of points of $\mathbb{R}^n$. If α is an inert value of P and $c_P(t) < t$ for all $t \in [\alpha, \beta]$, then there exists a sequence of collapses from $\mathcal{C}(P, \beta)$ to $\mathcal{C}(P, \alpha)$.*

4 Applications to 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 4.2). This requires to study in more details convexity defects functions, establishing connections with the distance function to X in Section 4.1 and the stability of h_X. Finally, we construct a bridge between shapes with an upper bounded convexity defects function and shapes with a positive μ-reach in Section 4.3. We then compute in Section 4.4 the lowest density of points authorized by our theorems for a correct reconstruction.

4.1 Characterizing critical points of the distance function

We begin by giving 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. For this, 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. Specifically, Grove defines in [Gro93, page 360] critical points for the distance function to a closed subset of a Riemannian manifold. Using Equation (1.1)' in [Gro93, page 360], 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 3. *We say that $y \in \mathbb{R}^N$ is a* critical point *of the distance function $d(\cdot, X)$ if $y \in \mathrm{Hull}(\Gamma_X(y))$. The* critical values *of $d(\cdot, X)$ are the images by $d(\cdot, X)$ of its critical points.*

Slightly recasting Proposition 1.8 in [Gro93, page 362], we have:

Theorem 3 (Isotopy Theorem [Gro93]). *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 value in the interval $[\alpha, \beta]$, then X^β deformation retracts to X^α.*

In section 3.1, we noted that $c_X(t) \leq h_X(t) \leq t$ for all t. Next lemma establishes that equality is attained if and only if t is a critical value of the distance function to X (see Figure 4).

Lemma 4. *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) $h_X(t) = t$.*

4.2 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 ε or less. Reconstruction results relie on the stability of h_X under perturbations of X (see Figure 4):

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

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 [AL10], 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 4, we deduce immediately that $X^{\alpha+\varepsilon}$ deformation retracts to P^α whenever the following two conditions are fulfilled:

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

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 t-offsets of X in the interval $[\alpha - 2\varepsilon, \alpha + 3\varepsilon]$ are all homotopy equivalent. We summarize our findings in the following theorem:

Theorem 4. *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 ε or less. The Čech complex $\mathcal{C}(P, \alpha)$ is homotopy equivalent to the ($\alpha - 2\varepsilon$)-offset of X whenever $h_X(t) < t - 3\varepsilon$ for all $t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$.*

Reconstruction with the Rips complex. If furthermore we suppose that $c_P(\vartheta_n\beta) < 2\alpha - \vartheta_n\beta$, we can apply Theorem 1 and deduce that (α, β)-almost 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 1 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 4 is also fulfilled. We deduce the following theorem:

Theorem 5. *Let ε, α and β be three non-negative real numbers such that $\alpha \leq \beta$ and $\eta = 2\alpha - \vartheta_n\beta - 2\varepsilon > 0$. Let P be a finite set of points whose Hausdorff distance to a compact subset X is ε or less. Then, any (α, β)-almost Rips complex of P is homotopy equivalent to the η-offset of X whenever α is an inert value of P and $h_X(\vartheta_n\beta + \varepsilon) < 2\alpha - \vartheta_n\beta - 2\varepsilon$.*

4.3 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 larger class of objects than shapes with a positive reach in [CCSL09]. Even though the distance function to X is not differentiable, it is possible to define a *generalized gradient function* $\nabla_X : \mathbb{R}^n \setminus X \to \mathbb{R}$ that coincides with the usual gradient at points where $d(\cdot, X)$ is differentiable and that vanishes precisely at points that are critical [CCSL09]. Specifically,

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

The critical function $\chi_X : \mathbb{R} \to \mathbb{R}$ is defined by $\chi_X(t) = \inf_{d(y,X)=t} \|\nabla(y)\|$. For $0 < \mu \leq 1$, authors in [CCSL09] define the μ-reach of X as $r_\mu(X) = \inf\{t > 0,\ \chi_X(t) < \mu\}$. The terminology comes from the fact that $r_1(X)$ coincides with the usual reach of X. Observe that $r_\mu(X) \geq R$ is equivalent to $\chi_X(t) \geq \mu$ for all $t \in [0, R)$. Our first lemma provides a lower bound on χ_X at t, assuming an upper bound on c_X at t.

Lemma 6. *For all compact set $X \subset \mathbb{R}^n$, all $0 \leq \mu \leq 1$ and all $t \geq 0$, the following implication holds:*

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

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 χ_X over the interval $[0, R)$. It extends a result in [AL10] and says intuitively that the convex hull of point set $\sigma \subset X$ cannot be too far away from a shape X, assuming σ can be enclosed in a ball of small radius t and X has a positive μ-reach.

Lemma 7. *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 $0 \leq t < R$. Then, for all $0 \leq t \leq R$, one has:*

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

The upper bound on h_X is an arc of ellipse which tends to an arc of parabola as $\mu \to 0$; see Figure 5, 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 [AL10]. Equivalently, the graph of h_X is below the circle with radius R and center $(0, R)$.

4.4 Reconstructing shapes with a positive μ-reach

Given a shape X whose μ-reach R is positive 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 α. Computations were realized using a computer algebra system and details are skipped. In Appendix D, we give all the details when $\mu = 1$, $R = 1$ and $n = +\infty$.

Reconstruction with the Čech complex. Combining Theorem 4 and Lemma 7, we obtain that the Čech complex $\mathcal{C}(P, \alpha)$ is homotopy equivalent to $X^{\alpha - 2\varepsilon}$ for all $\alpha \in (2\varepsilon, R - 3\varepsilon]$ whenever

$$\frac{1 + \mu(1-\mu) - \sqrt{1 - \mu(2-\mu)\left(\frac{t}{R}\right)^2}}{\mu(2-\mu)} R \quad < \quad t - 3\varepsilon, \qquad \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon].$$

Eliminating the square root, we can replace the above inequality by $H < 0$ where H is a polynomial of degree 2 in t. It follows that the above condition holds whenever the absolute difference between the two roots λ_1, λ_2 of H is greater than 2ε. The condition $|\lambda_2 - \lambda_1| > 2\varepsilon$ can be rewrote as the positivity of a polynomial of degree 2 in ε and holds whenever ε is smaller than the greatest root $\varepsilon^{\mathrm{cech}}(\mu)$ whose value is:

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

Interestingly, $\varepsilon^{\mathrm{cech}}(\mu)$ does not depend on the dimension n. Plotting the ratio $\frac{\varepsilon^{\mathrm{cech}}(\mu)}{R}$ as a function of μ (see Figure 3, left), 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 [CCSL09]. Still, for $\mu = 1$, we get $\varepsilon^{\mathrm{cech}}(1) = \frac{-3+\sqrt{22}}{13} \approx 0.13$ which is not as good as the value $3 - \sqrt{8} \approx 0.17$ obtained in [NSW08]

Reconstruction with the Rips complex. Combining Theorem 4 and Lemma 7, we get that the Rips complex $\mathcal{R}(P, \alpha)$ is homotopy equivalent to $X^{2\alpha - \vartheta_n \alpha - 2\varepsilon}$ for all $\alpha \in \left(\frac{2\varepsilon}{2-\vartheta_n}, \frac{R-\varepsilon}{\vartheta_n}\right]$ whenever

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

As before, we can eliminate the square root, replacing the above inequality by $H < 0$ where H is a polynomial of degree 2 in ε and α. Since we are looking for the greatest value of ε for which $H < 0$, we may assume that $\frac{\partial H}{\partial \alpha} = 0$. Plugging the value of α for which $\frac{\partial H}{\partial \alpha} = 0$ in H, we get a polynomial of degree 2 in ε whose greatest root $\varepsilon_n^{\mathrm{rips}}(\mu)$ gives the supremum of ε for which the above inequality holds. Plotting the ratio $\frac{\varepsilon_n^{\mathrm{rips}}(\mu)}{R}$ as a function of μ, we observe that the ratio is only positive on a subinterval $(\mu_n^*, 1]$ of $(0, 1]$; see Figure 3, left. Hence, we can only guarantee that Rips complexes provide a correct reconstruction for shapes with a positive μ-reach when $\mu > \mu_n^*$. In Figure 3, middle, we plotted μ_n^* as a function of n. μ_n^* increases with n and we were able to prove using a computer algebra system that μ_n^* tends to $\sqrt{2\sqrt{2} - 2} \approx 0.91$ as $n \to +\infty$. In Figure 3, right, we plotted $\frac{\varepsilon_n^{\mathrm{rips}}(1)}{R}$ as a function of n. For a fixed R, $\varepsilon_n^{\mathrm{rips}}(1)$ decreases with n and similarly, we proved that $\lim_{n \to +\infty} \varepsilon_n^{\mathrm{rips}}(1) = \frac{2\sqrt{2-\sqrt{2}} - \sqrt{2}}{2+\sqrt{2}} R \approx 0.034R$.

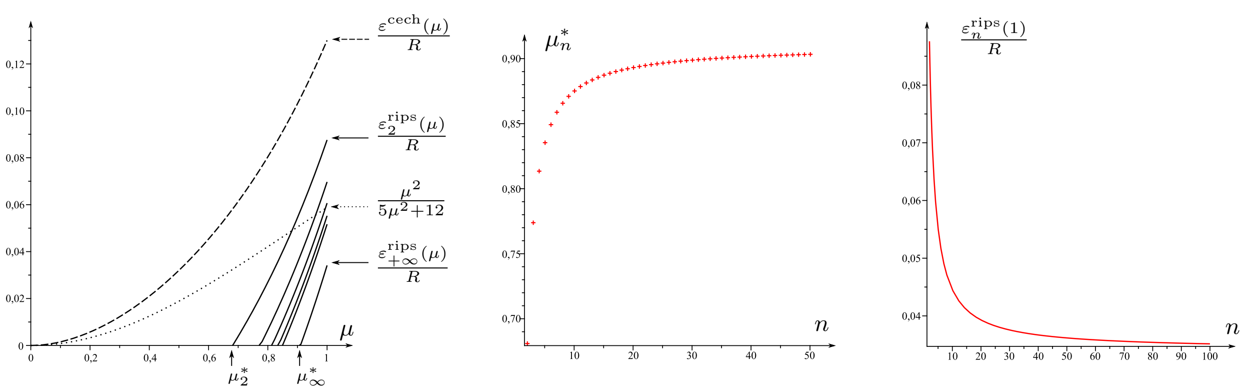

Figure 3: Left: Best ratios $\frac{\varepsilon}{R}$ we can get for a correct reconstruction of a shape with a positive μ-reach either with the Čech complex or the Rips complex for $n \in \{2, 3, 4, 5, 6, +\infty\}$; comparison with the ratio obtained in [CCSL09]. Middle: μ_n^* as a function of n. Right: $\frac{\varepsilon_n^{\mathrm{rips}}(1)}{R}$ as a function of n.

References

[AAD07] 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, Lousiana, USA, January 7–9 2007.

[AB99] N. Amenta and M. Bern. Surface reconstruction by Voronoi filtering. *Discrete and Computational Geometry*, 22(4):481–504, 1999.

[AB04] 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.

[ABE98] N. Amenta, M. Bern, and D. Eppstein. The crust and the β-skeleton: Combinatorial curve reconstruction. *Graphical Models and Image Processing*, 60(2):125–135, 1998.

[ABL03] 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.

[AL10] D. Attali and A. Lieutier. Reconstructing shapes with guarantees by unions of convex sets. In *26th Ann. Sympos. Comput. Geom.*, Snowbird, Utah, 2010. download, full-version.

[ALS] D Attali, A. Lieutier, and D. Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimension. Submitted to the 27th annual symposium on Computational Geometry (SoCG 2011).

[Att98] D. Attali. r-regular shape reconstruction from unorganized points. *Computational Geometry: Theory and Applications*, 10:239–247, 1998.

[BC02] J.D. 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.

[BDH09] J.D. 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.

[BG10] J.D. 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.

[Bjo96] A. Bjorner. Topological methods. In *Handbook of combinatorics (vol. 2)*, page 1850. MIT Press, 1996.

[BM09] J. A. Barmak and E. G. Minian. Strong homotopy types, nerves and collapses. Technical report, arXiv:0907.2954v1, 2009.

[CCSL09] 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.

[CdSEG10] E.W. Chambers, V. de Silva, J. Erickson, and R. Ghrist. Vietoris–rips complexes of planar point sets. *Discrete and Computational Geometry*, pages 1–16, 2010.

[CL05] F. Chazal and A. Lieutier. The λ-medial axis. *Graphical Models*, 67(4):304–331, 2005.

[CL08] 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.

[CO08] F. Chazal and S. Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proc. ACM Symposium on Computational Geometry*, pages 232–241, 2008.

[DEG99] T. K. Dey, H. Edelsbrunner, and S. Guha. Computational topology. In B. Chazelle, J. E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*, volume 223 of *Contemporary Mathematics*. AMS, Providence, 1999.

[Dey07] T.K. Dey. *Curve and surface reconstruction: algorithms with mathematical analysis*. Cambridge Univ Pr, 2007.

[DS08] V. De Silva. A weak characterisation of the Delaunay triangulation. *Geometriae Dedicata*, 135(1):39–64, 2008.

[dSG07] V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7:339–358, 2007.

[Ede93] 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.

[EM92] H. Edelsbrunner and E.P. M
"ucke. Three-dimensional alpha shapes. In *Proceedings of the 1992 workshop on Volume visualization*, pages 75–82. ACM, 1992.

[ES97] H. Edelsbrunner and N. Shah. Triangulating topological spaces. In *International Journal of Computational Geometry and Applications*, volume 7, pages 365–378. Citeseer, 1997.

[Gro93] K. Grove. Critical point theory for distance functions. In *Proc. of Symposia in Pure Mathematics*, volume 54, pages 357–386, 1993.

[Hau95] J.C. 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.

[Lat01] J. Latschev. Vietoris-Rips complexes of metric spaces near a closed Riemannian manifold. *Archiv der Mathematik*, 77(6):522–528, 2001.

[Lie04] A. Lieutier. Any open bounded subset of R n has the same homotopy type as its medial axis. *Computer-Aided Design*, 36(11):1029–1046, 2004.

[Mun93] J.R. Munkres. *Elements of algebraic topology*. Perseus Books, 1993.

[NSW08] 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.

[Wel91] 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.

[Zom10a] A. Zomorodian. Fast construction of the Vietoris-Rips complex. *Computers & Graphics*, 2010.

[Zom10b] A. Zomorodian. The tidy set: a minimal simplicial set for computing homology of clique complexes. In *Proceedings of ACM symposium of computational geometry*, 2010.

A Plotting Convexity Defects Functions

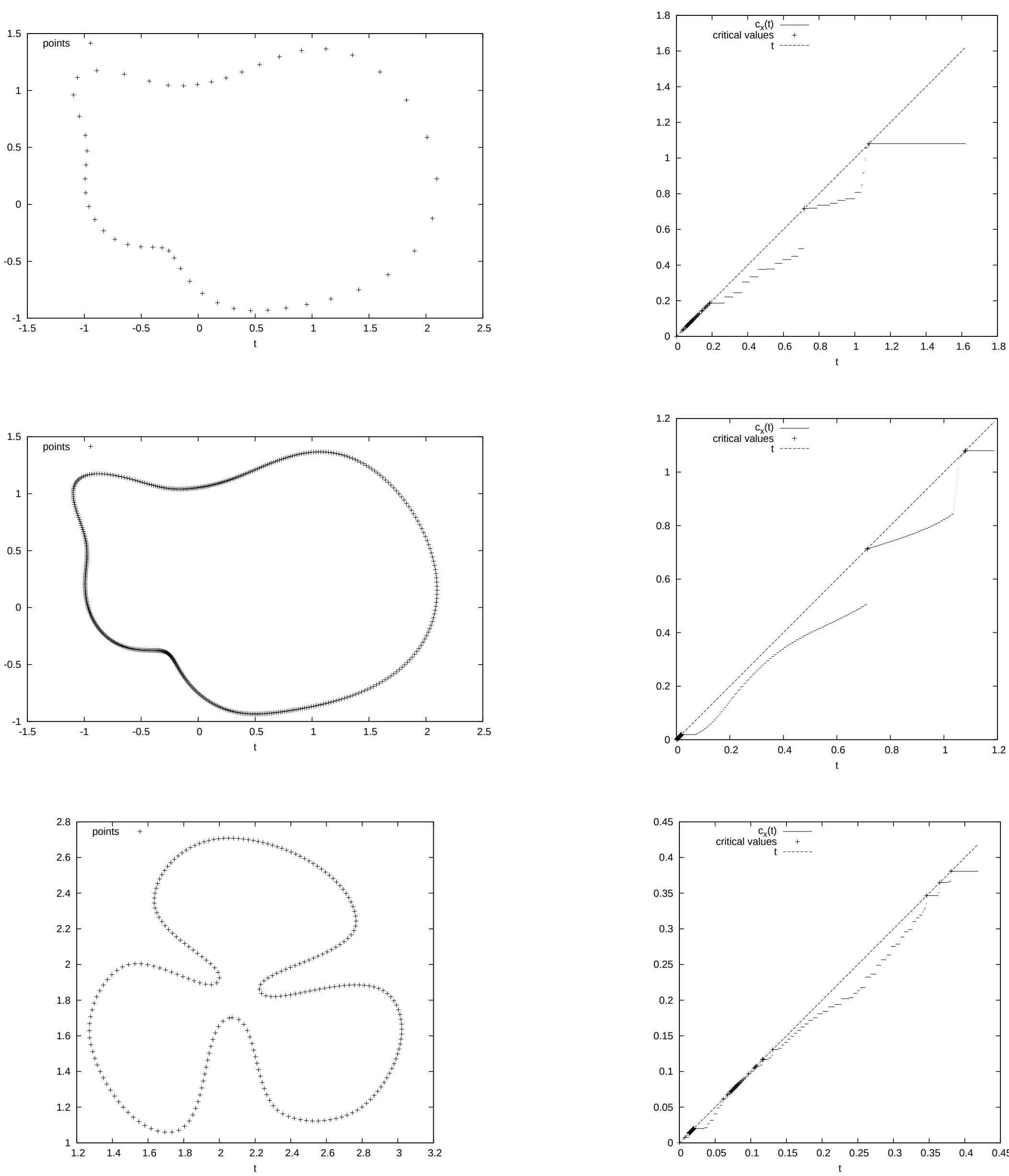

Figure 4: Various point sets P on the left and corresponding convexity defects function c_P on the right.

B Technical lemmas

Let us prove that the map $\sigma \mapsto \mathrm{Rad}(\sigma)$ is 1-Lipschitz:

Lemma 8. *For every non-empty bounded subsets σ and ξ of $\mathbb{R}^n$, we have $|\mathrm{Rad}(\sigma) - \mathrm{Rad}(\xi)| \leq d_H(\sigma, \xi)$.*

Proof. Writing B be for the smallest ball enclosing σ, we have $\xi \subset \sigma^\varepsilon \subset B^\varepsilon$, showing that $\mathrm{Rad}(\xi) \leq \mathrm{Rad}(\sigma) + \varepsilon$. $\square$

Given a point set $P \subset \mathbb{R}^n$, we say that a map $f : P \to \mathbb{R}^n$ is an *ε-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 9. *Let $P \subset \mathbb{R}^n$ be a finite set of points. Consider two real numbers $\beta \geq \alpha \geq 0$ such that*

$$c_P(\vartheta_{\mathrm{n}}\, \beta) < 2\alpha - \vartheta_{\mathrm{n}}\, \beta$$

and suppose moreover that none of the smallest balls enclosing subsets of P has a radius equal to α. Then, there exist $\varepsilon > 0$ and $\beta' > \beta$ such that for all ε-small perturbations f of P, we have:

(i) $c_{f(P)}(\vartheta_{\mathrm{n}}\, \beta') < 2\alpha - \vartheta_{\mathrm{n}}\, \beta'$;

(ii) $\mathcal{C}(f(P), \alpha) = f(\mathcal{C}(P, \alpha))$;

(iii) if $\mathrm{Flag}\, G$ is an (α, β)-almost Rips complex of P, then $\mathrm{Flag}\, f(G)$ is an (α, β')-almost Rips complex of $f(P)$.

Proof. Let us establish (i). For this, set $t = \vartheta_{\mathrm{n}}\, \beta$ and define $\bar{t} = \min\{\mathrm{Rad}(\sigma) \mid \emptyset \neq \sigma \subset P \text{ and } \mathrm{Rad}(\sigma) > t\}$. By construction, $\bar{t} > t$. Lemma 5 ensures that for all subsets $P' \subset \mathbb{R}^n$ within Hausdorff distance ε from P and for all $t' \geq 0$, the following implication holds:

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

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

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

and (i) is proved with $\beta' = t' / \vartheta_{\mathrm{n}}$. 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) $\mathrm{Rad}(\sigma) \notin [\alpha - \varepsilon, \alpha + \varepsilon]$ for all $\emptyset \neq \sigma \subset P$. Let f be an ε-small perturbation of P. Using Lemma 5 and condition (5), we get

$$\sigma \in \mathcal{C}(P, \alpha) \Leftrightarrow \mathrm{Rad}(\sigma) \leq \alpha \Leftrightarrow \mathrm{Rad}(\sigma) \leq \alpha - \varepsilon \Rightarrow \mathrm{Rad}(f(\sigma)) \leq \alpha \Leftrightarrow f(\sigma) \in \mathcal{C}(f(P), \alpha)$$

and

$$f(\sigma) \in \mathcal{C}(f(P), \alpha) \Leftrightarrow \mathrm{Rad}(f(\sigma)) \leq \alpha \Rightarrow \mathrm{Rad}(\sigma) \leq \alpha + \varepsilon \Leftrightarrow \mathrm{Rad}(\sigma) \leq \alpha \Leftrightarrow \sigma \in \mathcal{C}(P, \alpha),$$

yielding (ii). Consider a graph G whose flag complex is an (α, β)-almost 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). $\square$

C Missing proofs

Proof of Lemma 1. Let σ be a non-empty bounded set of $\mathbb{R}^n$. We first establish the existence of a smallest ball enclosing σ. 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

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

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 σ and whose radii are smaller than or equal to the diameter of σ is

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

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 [Wel91]. $\square$

Proof of Lemma 3. Suppose first that K contains a vertex v whose link is a cone with apex o. Slightly adapting the proof of Proposition 2.9 in [BM09], we prove that $|K|$ deformation retracts to $|K \setminus \mathrm{St}_K(v)|$. Define a vertex map $\pi : \mathrm{Vert}(K) \to \mathrm{Vert}(K)$ which is the identity on $\mathrm{Vert}(K) \setminus \{v\}$ and such that $\pi(v) = o$. If τ is a proper coface of v, then $\tau \setminus \{v\}$ belongs to the link of v and because the link is a cone with apex o, it also contains $\pi(\tau) = (\tau \setminus \{v\}) \cup \{o\}$. Moreover, $\pi(\tau) \cup \tau = \tau \cup \{o\}$ belongs to K. It follows that π can be extended to a simplicial map which is contiguous to the identity of K. Furthermore, $\pi(K) = K \setminus \mathrm{St}_K(v)$ and the restriction of π to $K \setminus \mathrm{St}_K(v)$ is the identity. Thus, the map $H : |K| \times [0,1] \to |K|$ defined by $H(x,t) = (1-t)x + t\pi(x)$ is a deformation retraction of $|K|$ onto $|K \setminus \mathrm{St}_K(v)|$.

Suppose now σ is a simplex in K whose link is a cone with apex o. We reduce this case to the previous one by subdividing simplices in the star of σ as follows. Let $\hat{\sigma}$ be the barycenter of σ and let $\mathrm{Bd}\,\sigma$ designate the set of proper faces of σ. We build a simplicial complex K' from K, replacing the closed star $\overline{\mathrm{St}}_K(\sigma)$ by the join $\{\hat{\sigma}\} * \mathrm{Bd}\,\sigma * \mathrm{Lk}_K(\sigma)$. Note that if σ is a vertex, then the join coincides with the closed star of σ and $K' = K$. By construction, the simplicial complex K and its subdivision K' have in common the set of simplices $K \setminus \mathrm{St}_K(\sigma) = K' \setminus \mathrm{St}_{K'}(\hat{\sigma})$. Let us show that the link of $\hat{\sigma}$ in K' is a cone with apex o. By construction, $\mathrm{Lk}_{K'}(\hat{\sigma}) = \mathrm{Bd}\,\sigma * \mathrm{Lk}_K(\sigma)$. Using the existence of a subcomplex $L \subset K$ such that $\mathrm{Lk}_K(\sigma) = \{o\} * L$, we get that $\mathrm{Lk}_{K'}(\hat{\sigma}) = \{o\} * \mathrm{Bd}\,\sigma * L$ is a cone. The first part of the proof implies that $|K| = |K'|$ deformation retracts to $|K \setminus \mathrm{St}_K(\sigma)| = |K' \setminus \mathrm{St}_{K'}(\hat{\sigma})|$. $\square$

Proof of Lemma 4. Making $x = y$ in Lemma 2, we observe that if $y \in \mathrm{Hull}(\sigma)$ satisfies $d(y,\sigma) \geq t$ and $\mathrm{Rad}(\sigma) \leq t$, then $y = \mathrm{Center}(\sigma)$.

Let us prove that (1) $\implies$ (2). Consider a critical point y whose distance to X is t. Setting $\sigma = \Gamma_X(y)$, we have $y \in \mathrm{Hull}(\sigma)$, $d(y,\sigma) = t$ and $\mathrm{Rad}(\sigma) \leq t$. Thanks to our observation, it follows that $y = \mathrm{Center}(\sigma)$ and consequently $c_X(t) = t$. Because $c_X(t) \leq h_X(t) \leq t$, we have (2) $\implies$ (3). 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 $\mathrm{Rad}(\sigma) \leq t$ and $y \in \mathrm{Hull}(\sigma)$ such that $d(y,\sigma) \geq d(y,X) = t$. Our observation then implies that $d(y, \mathrm{Center}(\sigma)) = 0$. Hence, $y = \mathrm{Center}(\sigma)$, $t = \mathrm{Rad}(\sigma)$ and σ represents a set of points in X with minimum distance to y. Since $y \in \mathrm{Hull}(\sigma) \subset \mathrm{Hull}(\Gamma_X(y))$, it follows that y is a critical point of the distance function to X, which concludes the proof. $\square$

Proof of Lemma 5. Consider a non-empty subset $\sigma \subset P$ with $\mathrm{Rad}(\sigma) \leq t$ and set $\xi = X \cap \sigma^{\varepsilon}$. By construction, ξ is non-empty and $d_H(\xi, \sigma) \leq \varepsilon$. Hence, Lemma 8 implies that $\mathrm{Rad}(\xi) \leq t + \varepsilon$. Using

$\mathrm{Hull}(\xi^\varepsilon) = \mathrm{Hull}(\xi)^\varepsilon$, we get that $\mathrm{Hull}(\sigma) \subset \mathrm{Hull}(\xi)^\varepsilon \subset X^{h_X(t+\varepsilon)+\varepsilon} \subset Y^{h_X(t+\varepsilon)+2\varepsilon}$, yielding the result. $\square$

Proof of Lemma 6. Consider $y \in \mathbb{R}^n$ whose distance to X is 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. Suppose the smallest ball enclosing σ has center z and radius s. Since $s \leq t$, we get $c_X(s) \leq c_X(t) < (1-\mu)t$ and therefore $t - \|y-z\| \leq d(z, X) < (1-\mu)t$. It follows that $\|\nabla_X(y)\| = \frac{\|z-y\|}{t} > \mu$. $\square$

Proof of Lemma 7. Given $\sigma \subset X$ with $\mathrm{Rad}(\sigma) \leq R$ and $y_0 \in \mathrm{Hull}(\sigma)$, we establish an upper bound on $d(y_0, X)$ expressed as a function of $\mathrm{Rad}(\sigma)$. Consider an integral line $\mathcal{C}_{y_0}$ of the flow associated to the distance function to X and starting at point y_0 [Lie04, CL05]. Suppose this integral line is parameterized by arc length and set $y_s = \mathcal{C}_{y_0}(s)$. For $s < R - d(y_0, X)$, one has $d(y_s, X) \leq d(y_0, X) + s < R$ and therefore $\chi_X(d(y_s, X)) \geq \mu$ which implies $\|\nabla_X(y_s)\| \geq \mu$. In particular, 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 $[0, R - d(y_0, X)]$. Since the norm of the gradient $\|\nabla_X(y_s)\| \geq \mu$ is equal to the right derivative of $s \mapsto d(y_s, X)$ (see [Lie04, CL05]), we obtain that

$$\frac{d(y_s, X) - d(y_0, X)}{s} \quad \geq \quad \mu.$$

Applying Lemma 2 with $x = y_s$ and $y = y_0$ gives $d(y_s, X)^2 \leq d(y_s, \sigma)^2 \leq s^2 + \mathrm{Rad}(\sigma)^2$ from which we deduce the inequality $(d(y_0, X) + \mu s)^2 \leq s^2 + \mathrm{Rad}(\sigma)^2$. Plugging $s = R - d(y_0, X)$, setting $\delta = \frac{d(y_0,X)}{R}$, $\rho = \frac{\mathrm{Rad}(\sigma)}{R}$ and rearranging this inequality gives us

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

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. $\square$

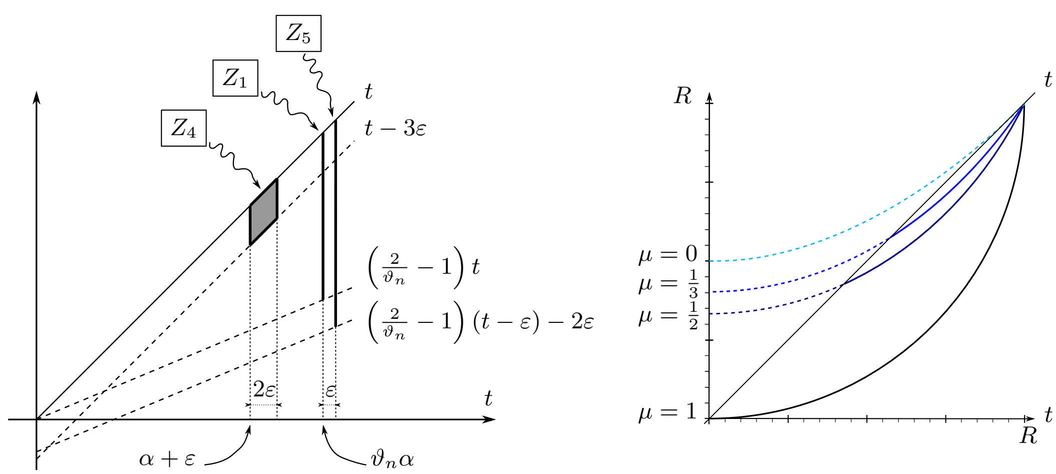

Figure 5: Left: For $i \in \{1, 4, 5\}$, the hypotheses of Theorem i are depicted as regions Z_i avoided by the graph of a convexity defects function. Specifically, if $c_P \cap Z_1 = \emptyset$, Theorem 1 implies $\mathcal{R}(P, \alpha) \simeq P^\alpha$. If $h_X \cap Z_4 = \emptyset$, Theorem 4 implies $\mathcal{C}(P, \alpha) \simeq X^{\alpha-2\varepsilon}$. If $h_X \cap Z_5 = \emptyset$, Theorem 5 implies $\mathcal{R}(P, \alpha) \simeq X^{2\alpha-\vartheta_n\alpha-2\varepsilon}$. Right: upper bounds on h_X in Lemma 7 for $\mu \in \{0, \frac{1}{3}, \frac{1}{2}, 1\}$.

D Reconstructing shapes with a positive reach

In this appendix, we redo computations of Section 4.4, setting $\mu = 1$, $R = 1$, $n = +\infty$.

Reconstruction with the Čech complex. Combining Theorem 4 and Lemma 7, we get that the Čech complex $\mathcal{C}(P,\alpha)$ is homotopy equivalent to $X^{\alpha-2\varepsilon}$ for $\alpha \in (2\varepsilon, 1-3\varepsilon]$ whenever

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

which can be rewrote as

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

This condition holds whenever the absolute difference between the two roots of the polynomial in t is greater than 2ε, *i.e.* whenever $0 > 13\varepsilon^2 + 6\varepsilon - 1$. The supremum of ε for which the previous equation holds is $\varepsilon^{\mathrm{cech}}(1) = \frac{-3+\sqrt{22}}{13} \approx 0.13$.

Reconstruction with the Rips complex. Combining Theorem 4 and Lemma 7, we get that the Rips complex $\mathcal{R}(P,\alpha)$ is homotopy equivalent to $X^{2\alpha-\sqrt{2}\alpha-2\varepsilon}$ for all $\alpha \in \left(\frac{2\varepsilon}{2-\sqrt{2}}, \frac{1-\varepsilon}{\sqrt{2}}\right]$ whenever

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

which we can rewrote as

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

Since we are looking for the greatest value of ε for which the above equation holds, we may assume that he partial derivative of the left side with respect to α 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

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

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