Equivalent Euclidean Data Complexes

Kristian André Jakobsen

June 2021

Master's Thesis in Topology

Department of Mathematics
University of Bergen

Abstract

Euclidean data complexes are simplicial complexes that have been constructed from a point cloud in Euclidean space. Two of the most important examples of such complexes are the Čech and Alpha complex. In this thesis, we will prove that these are homotopy equivalent to the Delaunay-Čech complex using the geometric and gradient collapse arguments. Moreover, we introduce a new Euclidean data complex that we call the selective Delaunay-Alpha complex. Not only does it generalize the other three, but it is also simple-homotopy equivalent to them. The implications of this result will also be discussed.

Contents

Acknowledgements

I have several people I should thank for helping me get this thesis done. First and foremost, I am very grateful to my advisor, Morten Brun, for suggesting an exciting and challenging project. He has patiently provided help whenever I have been stuck in things related to mathematics or career. I really appreciate doing mathematics together because it always feels like a process of discovery for both. I have learned a lot the past two years thanks in no small part to Morten.

I also need to thank my friends and family. I want to thank Erlend Raa Vågset for reading through my thesis and providing some useful pointers. Peter Brosten has not only been an amazing and hilarious flatmate, but also someone who has always been willing to proofread my thesis when few others could. I will miss our fun and miserable times doing math on the board. Alessandra Oshiro has also been incredibly supportive throughout the last year. Even if she couldn't check my proofs, Ale has always been eager to help in any other capacity. Last, but not least, thanks to the Jakobsens: Oscar, Svanhild and Kenneth. I couldn't ask for a more kind and loving family.

1 Introduction

Topological data analysis (TDA) is a subfield of mathematics that studies the shape of data using topological techniques. It lies at the intersection of computational geometry and topology, algebraic topology, and data analysis. To help make the problem of finding useful patterns tractable, we usually make some general assumptions about the kind of data we study. First, we assume it consists of a finite set $X \subseteq \mathbb{R}^d$. Perhaps somewhat surprisingly, this does not limit the scope of the kind of problems we can tackle. Real-world data – whether it is geospatial, times series or image-based – can be represented as a collection of finite vectors in some d-dimensional Euclidean space. This can be good to keep in mind whenever TDA seems overly abstract.

Second, the *manifold hypothesis* is the assumption that these points have been sampled from a lower dimensional submanifold of $\mathbb{R}^d$, e.g., a surface in $\mathbb{R}^3$, or a curve in $\mathbb{R}^2$. Nature is full of regularities and patterns that one would expect to be observed in the data as a whole. Hence, it makes theoretical sense to assume that we are not dealing with a uniform distribution in $\mathbb{R}^d$, and that the data varies smoothly because there are correlations between nearby data points. The manifold hypothesis is important for machine learning models to be able to extract features from the dataset with appropriate representations [4, p. 173]. The mathematical tools from TDA make the study of such manifolds interesting in their own right.

The manifold hypothesis has long been a theoretical assumption in the machine learning community, but TDA has provided tools to also justify it empirically [23]. *Persistent homology* is one such tool. It lets us distinguish between the holes and components that are in the manifold and those that are artefacts of noisy data [7]. As a result, we might try to classify a manifold by the homological features that persist as one builds an increasingly larger space from the data points in X. As a real-world example, Gunnar Carlsson et al. [8] used persistent homology to show that the space of 3×3 patches of grayscale natural images has the shape of a Klein bottle. Persistent homology is useful for data where connectivity and loops matter, with applications in viral evolution [9] and remote sensing [14].

However, in order to compute persistent homology, we need to have a family of *simplicial complexes* that are ordered by inclusion, called a *filtration*. Hence, this thesis is concerned with the step before one computes persistent homology. The focus is on ways to reconstruct, or approximate, the underlying manifold based on data that was sampled from it, and produce a filtration of simplicial complexes [21, p. 28]. Note that we do not concern ourselves with how good these approximations are. We should rather use it as motivation for why these complexes might be interesting. We will construct different kinds of simplicial complexes, but all of them have X as their set of vertices. Figure 1 is a geometric

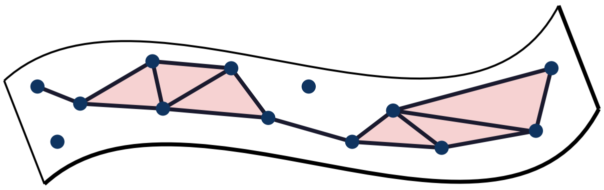

Figure 1: The data points are used to construct a simplicial complex that approximates the manifold that the data is sampled from.

illustration of how such a reconstruction might look. An *n-simplex* between $n+1$ points in X can be thought of as an approximation of the manifold in the region of those points, given they are sufficiently close. Better yet, the dimension of the simplex formed can tell us something about the density of the data in said region. This has obvious parallels to standard *clustering* methods. We will call simplicial complexes that have been constructed from a finite set of points in $\mathbb{R}^d$ *Euclidean data complexes*. This is not to be confused with the concept of *data complexes* by Abraham D. Smith et al. [1]. Euclidean data complexes not considered in this thesis include the *Vietoris-Rips*, *Witness* and *Intrinsic Čech complex*.

However, different measures of "nearness" produce simplices of differing dimensions and computational complexity. One of the more obvious approaches in constructing a Euclidean data complex would be to grow balls around each point in X with some radius $r \geq 0$ to form a cover that locally resembles the underlying manifold. Then we can form an n-simplex if the intersection of $n+1$ balls is non-empty. This is known as the *Čech complex*, and, by the *Nerve Theorem*, it is homotopy equivalent to the union of the balls [21, p. 31]. One problem with this Euclidean data complex is how quickly high dimensional simplices are formed. If the number of points in X exceeds the dimension of the ambient Euclidean space, then a sufficiently large r will create simplices that cannot be embedded in the ambient space. Not only is this computationally inefficient, but it also goes against one of the motivations for studying these complexes in the first place. The *Alpha complex* attempts to resolve these issues by limiting the number of balls that can intersect, with only very specific configurations of the data producing high dimensional simplices. The goal of this thesis is to show that these two Euclidean data complexes are *homotopy equivalent* to a complex that includes the Alpha complex and is included by the Čech complex, namely the *Delaunay-Čech complex*. In other words, we do not lose much topological information by dealing with either the Alpha, Čech or Delaunay-Čech complex. There will be given two different arguments for this result. One will be called the *geometric argument*, and the other the *gradient collapse argument*. These were given by Nello Blaser and Morten Brun [5], and Ulrich Bauer and Herbert Edelsbrunner [3], respectively. In the process, we will prove a more general result for a new simplicial complex that generalizes the ones mentioned above.

The thesis is structured as follows:

Section 2 will introduce the two Euclidean data complexes mentioned above, and others that play an important role. It will also show how they relate to their respective *radius functions*. The main contribution in this section is the *selective Delaunay-Alpha complex*.[1] It generalizes all the other Euclidean data complexes in this section.

Section 3 will show an alternative way of understanding the Euclidean data complexes introduced in Section 2 by using *dissimilarities*. The most important result in this section is the *Dowker Nerve Theorem* since it will be the basis for the geometric argument. The main contribution in this section is defining Euclidean data complexes, and proving several smaller results, that are not mentioned in [5]. This helps bridge the gap to the approach in [3].

Section 4 will introduce the discrete version of *Morse theory*. The techniques here will be essential for the gradient collapse argument. The main contribution in this section is relating definitions in [3] to those in [18]. We also provide proofs of Theorem 4.12 and Theorem 4.16 that were omitted in [3]. The most important result will be showing that the radius functions from Section 2 are *generalized discrete Morse functions* by adding a small assumption to our dataset.

Section 5 will present the two arguments for why the Alpha and Čech complexes are equivalent to the Delaunay-Čech complex. The main contribution in this section is formulating results in [3] in terms of the selective Delaunay-Alpha complex, which leads to corollary that proves this complex is simple-homotopy equivalent to all the complexes it generalizes.

Section 6 will summarize the results in this thesis and discuss an unsuccessful approach in giving a geometric proof of showing that the selective Delaunay-Alpha and Čech complexes are homotopic. It will also discuss some implications of having equivalent Euclidean data complexes.

[1] Note that this complex is implicit in Theorem 5.9 in [3], but was independently discovered using dissimilarities.

2 Euclidean Data Complexes

Before explaining what we mean by Euclidean data complexes, we need to define simplicial complexes. There are two notions of simplicial complexes: geometric and abstract. *Geometric simplicial complexes* are topological spaces and intuitive, but can be cumbersome to work with. *Abstract simplicial complexes*, on the other hand, contain the necessary combinatorial information for computations, but do not have the topological properties that a geometric simplicial complex has. This thesis will exclusively work with *abstract* simplicial complexes. Hence, it should be assumed that it is the latter notion we refer to when talking about "simplicial complexes", unless stated otherwise. We begin by defining the central objects of our study.

2.1 Abstract Simplicial Complexes

Those familiar with geometric simplicial complexes will recognize that abstract simplicial complexes are also built out of smaller simplices. One cannot, e.g., have a 2-simplex without also having three 1-simplices and three 0-simplices.

Definition 2.1. An *(abstract) simplicial complex* K on a set V is a collection of subsets of V such that if $\sigma \in K$ and $\tau \subseteq \sigma$ is non-empty, then $\tau \in K$.

We call V the *vertex set* of K and write $V(K)$ for this set. An element $\sigma \in K$ is called a *simplex* in K. Given another simplex $\tau \subseteq \sigma$ in K, then σ is a *coface* of τ, while τ is a *face* of σ. If the inclusion is proper, we say that σ is a *proper coface* and τ is a *proper face*. If $L \subseteq K$ is an inclusion of simplicial complexes, then L is said to be a *subcomplex* of K. The dimension p of a simplex σ is given by $|\sigma| - 1$ and we say that σ is a *p-simplex*. This is similar to the notion of dimension for a geometric simplex being the number of vertices minus one. The *dimension* of a simplicial complex is given by the maximum dimension of any of its simplices. We will only need to consider simplicial complexes defined on a finite vertex set V.

It is worth reiterating the distinction between the two types of simplicial complexes for those more familiar with the geometric version. In Figure 2 we have drawn a geometric illustration of an abstract simplicial complex. We can think of the vertices as being sampled from $\mathbb{R}^3$ and a p-simplex being drawn between $p+1$ points, if they form a simplex. In this case, we start out with 10 vertices and form a 3-simplex, a 2-simplex and two 1-simplices, if we do not count faces. Unlike geometric simplicial complexes, we only care about the relationship between simplices, regardless of potential intersections. We also want to talk about structure-preserving maps between simplicial complexes.

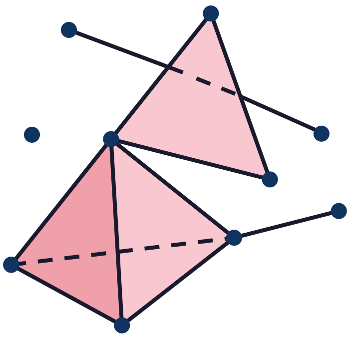

Figure 2: Geometric illustration of an abstract simplicial complex.

Definition 2.2. A *simplicial map* $f\colon K \to L$ between simplicial complexes K and L is defined as a function $f\colon V(K) \to V(L)$ such that a simplex $\sigma \in K$ gets mapped to a simplex $f(\sigma) \in L$.

There is a category **ASC** with abstract simplicial complexes as objects and simplicial maps as morphisms. It is useful to consider this category when we introduce other kinds of structures and want to relate them by means of functors. However, we will manage without categorical language. Now we make an abstract simplicial complex into a topological space.

Definition 2.3. The *geometric realization* $|K|$ of a simplicial complex K, with vertex set V, is the subspace of the function space $[0,1]^V$, given the strong topology[2], such that α is in $|K|$ if

(1) $\alpha^{-1}((0,1])$ is a simplex in K, and

(2) $\sum_{v \in V} \alpha(v) = 1$.

This abstract definition is more easily understood by considering each function α in $|K|$ as describing a convex combination of the vertices in a simplex $\sigma \in K$. The convex hull, i.e., the set of all convex combinations, represents the geometric realization $|\sigma|$ of σ. Although this is similar to how geometric simplicial complexes are defined [15], note that the geometric realization is not turning an abstract simplicial complex into a geometric one. The first condition ensures we only consider convex combinations of vertices that make up a simplex $\alpha^{-1}((0,1])$ that exists in K. The second tells us that α is a "point" in the simplex $\alpha^{-1}((0,1])$, defined by the vertices that become non-zero. Since every vertex $v_i \in V$ is itself a

[2]The *strong topology* is convenient when we want to look at functions into $|K|$, such as in the case of the Dowker Nerve Theorem [12, p. 355].

simplex, there must necessarily be a function $\pi_i \in |K|$ such that

$$\pi_i(v) = \begin{cases} 1 & \text{if } v = v_i \\ 0 & \text{otherwise.} \end{cases}$$

For a higher dimensional simplex σ, it will be as if we introduce a new axis for each vertex in σ and then find the points lying in-between them. Figure 3 illustrates this for a 1-simplex $\sigma = \{v_1, v_2\}$ and how $|\sigma|$ is the closure of

$$\{\alpha \colon V \to [0,1] \mid \alpha(v) > 0 \text{ for all } v \in \sigma\}.$$

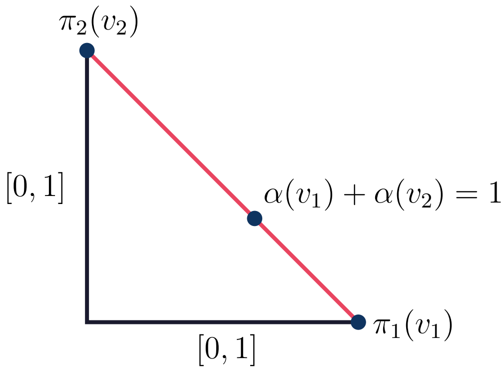

Figure 3: A function $\alpha \colon V \to [0,1]$ in $|K|$ considered as a point on a 1-simplex. The red line represents all such points, giving $|\{v_1, v_2\}| \subseteq |K|$.

We are now ready to state in what sense simplicial complexes in this thesis will be considered equivalent. It is, namely, in terms of their geometric realizations that simplicial complexes are equivalent.

Definition 2.4. Simplicial complexes K and L are *homotopy equivalent*, denoted $K \simeq L$, if there is a homotopy equivalence $|K| \simeq |L|$.

Since we will introduce a parameter for our complexes, it is useful to consider a family of simplicial complexes. We define the order on this family to be inclusion of subcomplexes.

Definition 2.5. A *filtration* of a simplicial complex K is a sequence of subcomplexes

$$\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K.$$ [3]

A simplicial complex is said to be *filtered* if it has a filtration.

[3] It is not strictly necessary to require that the sequence begins with the empty set.

Note that since we are working with finite complexes, any infinite sequence of subcomplexes will contain a finite subsequence such that each inclusion is proper. This is worth keeping in mind when we will be working with a parameter in $\mathbb{R}$.

2.2 Examples of Euclidean Data Complexes

As stated in the introduction, we are interested in how we can construct simplicial complexes from a finite set of data points – or a *point cloud* – X in $\mathbb{R}^d$. In this thesis, we call such simplicial complexes for *Euclidean data complexes*. However, not only are there several complexes one may construct from X, there are also two different ways of constructing the same Euclidean data complex. One is to directly give a condition on balls centered at the different points in the point cloud. The other uses continuous functions to define how dissimilar every point in $\mathbb{R}^d$ is from every point in X.

We begin, however, with the former way of defining the *Čech*, *Delaunay* and *Alpha complexes*, as they require less machinery. We will then see how they are particular instances of the *selective Alpha complex* and how they may be combined to form the *Delaunay-Čech complex*. Moreover, the Čech and Delaunay complexes can be described by smallest enclosing balls and smallest circumscribing spheres, respectively. This idea will also be generalized to the selective Alpha complex via *radius functions*. Finally, we introduce a new simplicial complex, called the *selective Delaunay-Alpha complex*, as a generalization of all the others. Besides these, there are several other Euclidean data complexes. Two of the most popular for computations are the *Vietoris-Rips* and *Witness complexes* [21].

Let $d(x, y)$ be the Euclidean distance between two points x and y in $\mathbb{R}^d$. We also define the *open ball* of radius $r \geq 0$ centered at x as

$$B_d(x, r) = \{y \in \mathbb{R}^d \mid d(x, y) < r\},$$

and denote its closure $\overline{B_d(x, r)}$ by $\bar{B}_d(x, r)$. The subscript indicates that one can use other metrics than d. We will revisit this point when considering dissimilarities.

2.2.1 Čech Complexes

Before defining the Čech complex, it can be informative to think in concrete terms about what property of a given dataset one is hoping to capture. When doing data analysis, one tries to find some useful pattern in the data. A reasonable hypothesis could be that a notion of nearness tells us when certain data points should be considered related. For instance, consider a dataset consisting of all the RGB values of every pixel in a collection of images. A clustering algorithm might identify different classes for the objects in the images by grouping together the images with similar RGB values in pixels that are sufficiently near each other. Images of lawns might thus be clustered together because they all have high green values in pixels near the bottom of the image. Geometrically, one could think of adding a p-simplex whenever $p+1$ data points are sufficiently close, and a higher value of p indicates a more dense clustering. It is this intuitive notion of "relatedness" that our simplicial complexes, including the Čech complex, are trying to extract from the point cloud. We will be using the Euclidean distance d to measure nearness, but there is no reason why one cannot use any other distance metric.

Definition 2.6. The *Čech complex* of a finite set $X \subseteq \mathbb{R}^d$ with radius $r \geq 0$ is the simplicial complex

$$\check{\mathrm{C}}\mathrm{ech}_r(X) = \left\{ \sigma \subseteq X \mid \bigcap_{x \in \sigma} \bar{B}_d(x, r) \neq \emptyset \right\}.$$

That is, we construct a p-simplex whenever $p+1$ closed balls, centered at $p+1$ different points in X, have a common intersection. Note that in the example of Figure 4 there are two triangles, but only one forms the boundary of a 2-simplex. Keep in mind that we have merely drawn a geometric simplicial complex as an

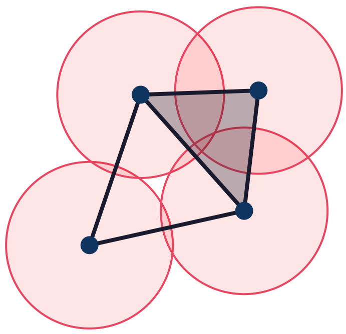

Figure 4: A geometric illustration of a Čech complex.

illustration of the combinatorial information in an abstract simplicial complex. The Čech complex gives a filtration of the *full simplex* $\Delta(X) = \mathcal{P}(X) - \{\emptyset\}$ spanned by X since $\check{\mathrm{C}}\mathrm{ech}_0(X)$ is empty and $\check{\mathrm{C}}\mathrm{ech}_\infty(X)$ is the full simplex, and since $r \leq s$ implies $\check{\mathrm{C}}\mathrm{ech}_r(X) \subseteq \check{\mathrm{C}}\mathrm{ech}_s(X)$. This filtration gets constructed by increasing r, adding simplices to the vertices in X until $\Delta(X)$ is formed.

Going back to our concrete example, one can think of the radius r of the closed balls as being the tolerance level for some data points to be considered "near". If $r = 0$, one would not tolerate any "noise" in the data, while increasing r indefinitely, one ends up considering all points to be "near" each other. An approach that looks at which topological features persist as one increases r is called *persistent homology* [7].

Implementing the definition, as given, is computationally inefficient. To avoid having to check all the intersections, we can give a description of the Čech complex that is faster. We reformulate the condition for a p-simplex to be added.

Proposition 2.7. *A set $\sigma \subseteq X$ containing $p+1$ points is a p-simplex in $\check{\mathrm{C}}\mathrm{ech}_r(X)$ if and only if there exists a point $q \in \mathbb{R}^d$ such that $\sigma \subseteq \bar{B}_d(q, r)$.*

Proof. Suppose $\sigma \subseteq \check{\mathrm{C}}\mathrm{ech}_r(X)$ is a p-simplex. Then we know there exists a $q \in \mathbb{R}^d$ such that q is in the intersection of the $p+1$ closed balls of radius r and centered at the points in σ. Then $d(q, x) \leq r$ for every $x \in \sigma$, proving one direction.

Conversely, we assume σ is contained in $\bar{B}_d(q, r)$ for some $q \in \mathbb{R}^d$. In other words, $d(q, x) \leq r$ for every $x \in \sigma$. This implies that the common intersection of all balls $\bar{B}_d(x, r)$, with $x \in \sigma$, must at least contain q, and hence be non-empty. $\square$

Using this new condition, we can write an algorithm that decides whether $\sigma \subseteq X$ is a simplex in $\check{\mathrm{C}}\mathrm{ech}_r(X)$. Define the *miniball* of σ to be the smallest closed d-ball that contains all points in σ. It can be found using Welzl's miniball algorithm [25]. After finding the miniball of σ, one checks if its radius is smaller than or equal to r. If it is, then we know that σ is a simplex in $\check{\mathrm{C}}\mathrm{ech}_r(X)$ by Proposition 2.7.

2.2.2 Delaunay Complexes

A simplicial complex that has more interesting properties than the Čech complex is the Delaunay complex, and its parameterized version, the Alpha complex. What is important from the point of view of computation is that the Alpha complex captures the same structure as the Čech complex, but is much more computationally efficient because it substantially limits the dimension of its simplices. As has become a reoccurring theme so far, there are several ways of constructing the Delaunay complex, and hence also the Alpha complex. The most intuitive requires a cover of $\mathbb{R}^d$ associated to X that is called the *Voronoi diagram* of X, which is made up of the *Voronoi cells* of each point $x \in X$. The reader curious to learn more about the Voronoi diagram – and its uses in everything from biology to engineering – should seek out [2].

Definition 2.8. Given a finite set $X \subseteq \mathbb{R}^d$, the *Voronoi cell* of $x \in X$ is the set

$$\mathrm{Vor}(X, x) = \{p \in \mathbb{R}^d \mid d(x, p) \leq d(y, p) \text{ for all } y \in X\}.$$

The Voronoi cell of x gives us the set of all points in $\mathbb{R}^d$ that are closer to x than any other point in X. One way to find $\mathrm{Vor}(X, x)$ is to first draw a hyperplane

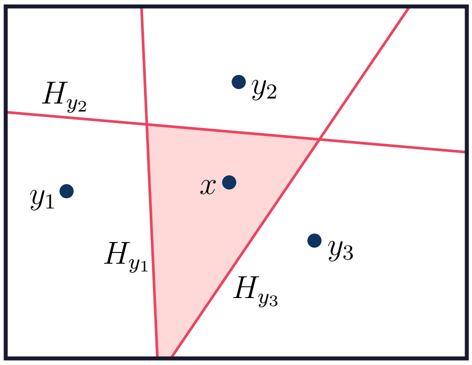

Figure 5: The Voronoi cell of x is the shaded area, and has been constructed from hyperplanes separating x from the other points in X.

H_y for every point $y \in X - \{x\}$ in such a way that it is equidistant to x and y. Let

$$H_y^+ = \{p \in \mathbb{R}^d \mid d(x,p) \leq d(y,p)\},$$

i.e., H_y^+ contains all points in $\mathbb{R}^d$ that are closer to x than y, or lie on H_y. Then we finally get that

$$\mathrm{Vor}(X,x) = \bigcap_{y \in X - \{x\}} H_y^+.$$

See Figure 5 for an example of how one would construct $\mathrm{Vor}(X,x)$ in the plane. The Voronoi cells of the points in X now form a cover $\{\mathrm{Vor}(X,x)\}_{x\in X}$ of $\mathbb{R}^d$. This cover is what we call the *Voronoi diagram* of X and denote it by $\mathrm{Vor}(X)$. Using the same points as in Figure 5, we get the Voronoi diagram on the left in Figure 6.

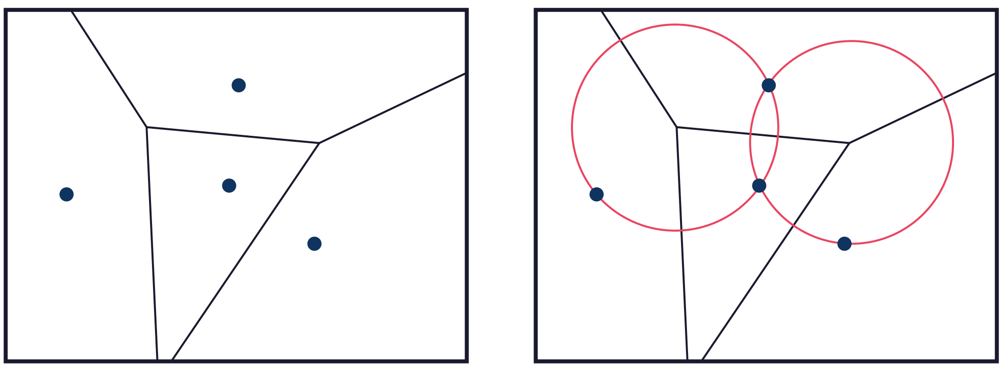

Figure 6: A Voronoi diagram for four points in $\mathbb{R}^2$ on the left. On the right, we have superimposed the empty circumscribed circles defining two 2-simplices.

Note that some of the Voronoi cells in $\mathrm{Vor}(X)$ intersect. This suggests a natural way to define a simplicial complex.

Definition 2.9. The *Delaunay complex* of a finite set $X \subseteq \mathbb{R}^d$ is the simplicial complex

$$\mathrm{Del}(X) = \left\{ \sigma \subseteq X \mid \bigcap_{x \in \sigma} \mathrm{Vor}(X, x) \neq \emptyset \right\}.$$

Before proving that the Delaunay complex decreases the dimension of its simplices when compared to the Čech complex, we need to qualify our claim first. We can, namely, have a special case in which all n points in our point cloud $X \subseteq \mathbb{R}^d$ lie on a common sphere. It is clear that their Voronoi cells would then intersect at the center of said sphere. In this special case, $\mathrm{Del}(X) = \check{\mathrm{C}}\mathrm{ech}_r(X)$ as r goes to ∞, irrespective of n. Hence, we will need to restrict the number of points that can lie on a common sphere. A good limiting number may be $d + 1$ since this will have the added benefit of being able to geometrically represent the Delaunay complex as embedded in $\mathbb{R}^d$.

Proposition 2.10. *If at most $d+1$ points in a finite set $X \subseteq \mathbb{R}^d$ lie on a common $(d-1)$-sphere, then the dimension of any simplex in* $\mathrm{Del}(X)$ *is at most d.*

Proof. Recall that one cannot have a simplex of dimension larger than $d+1$ unless it contains a $(d+1)$-simplex. Hence, it is sufficient to show that $\mathrm{Del}(X)$ does not contain a $(d+1)$-simplex whenever the points in X do not lie on a common sphere. Suppose $\mathrm{Del}(X)$ contains a $(d+1)$-simplex σ. Then the common intersection of the Voronoi cells of the $d + 2$ vertices in σ is non-empty. By definition of $\mathrm{Vor}(X, x)$, that means there is a point in $\mathbb{R}^d$ that is equidistant from every vertex in σ, defining a common $(d-1)$-sphere for $d + 2$ points in X. The result follows by contradiction. □

This naturally leads to the following description of the Delaunay complex that is reminiscent of the miniballs for Čech complexes. Instead of finding the smallest ball containing points in σ, we need to find the smallest *empty circumscribed $(d-1)$-sphere* of σ. In other words, we are looking for a $(d-1)$-sphere S such that all points in σ lie on S and no points in $X - \sigma$ are on, or bounded by, S. See Figure 7 for an example and a non-example. The smallest such sphere is called the *Delaunay sphere* of σ, and σ must be a simplex in $\mathrm{Del}(X)$. On the right of Figure 6, we get two empty circumscribed circles of three points each. These define two 2-simplices. Their center lies where the Voronoi diagrams of the respective points meet.

Note that a $(d-1)$-sphere is uniquely determined by $d+1$ *affinely independent* points. Hence, if a sphere circumscribes $d+1$ points in X, but still contains other points of X in its closure, then we know the former points cannot make up a simplex

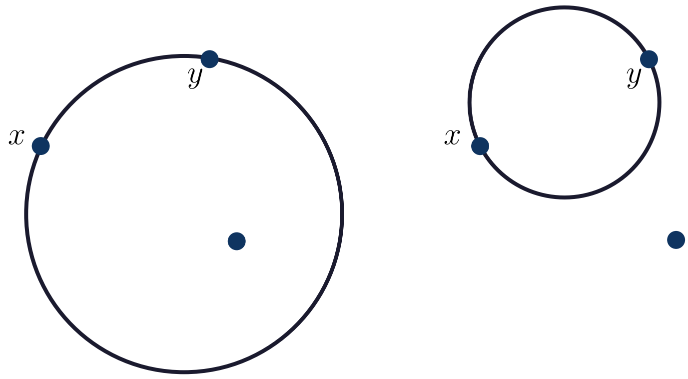

Figure 7: Although the circumscribed circle of x and y on the left is not empty, the one on the right is. So $\{x, y\}$ would be considered a simplex in $\mathrm{Del}(X)$

in $\mathrm{Del}(X)$. In the case of n affinely independent points, where $n \leq d$, there will exist many circumscribed spheres, but the smallest one is unique. Moreover, a lemma due to Seidel says that the center of this ball must lie in the convex hull of all n points [17]. It will be convenient to refer to the case in which we are guaranteed the existence of a circumscribed sphere for any n-simplex in $\mathrm{Del}(X)$ for $n \leq d$. Furthermore, we want this sphere not to contain any other points than those in σ.

Definition 2.11. A finite set $X \subseteq \mathbb{R}^d$ is in *general position* if for every subset $\sigma \subseteq X$, of at most $d+1$ points,

(1) σ is affinely independent, and

(2) $X - \sigma$ does not contain a point on the smallest circumscribed sphere of σ.

It can be hard to determine when there exists a circumscribed sphere in the case of $n > d+1$ points. We note, however, that the vertices of the platonic solids always have a circumscribed sphere. More generally, the center of the smallest circumscribed sphere of points in general position must lie in points' affine hull.

2.2.3 Alpha Complexes

Unlike the Čech complex, the Delaunay complex is a fixed complex. However, one can make it a filtered simplicial complex by letting it be a Čech complex that is bounded by the Delaunay complex. More formally, let

$$\mathrm{Vor}_r(X, x) = \mathrm{Vor}(X, x) \cap \bar{B}_d(x, r)$$

be the *Voronoi ball* of x with radius r.

Definition 2.12. The *Alpha complex* of a finite set $X \subseteq \mathbb{R}^d$ with radius $r \geq 0$ is the simplicial complex

$$\mathrm{Alpha}_r(X) = \left\{ \sigma \subseteq X \mid \bigcap_{x \in \sigma} \mathrm{Vor}_r(X, x) \neq \emptyset \right\}.$$

As one grows a closed ball, it will eventually meet the boundary of two or more Voronoi cells and stop growing in that direction. The result is akin to buns rising in the oven and getting stuck to each other, as seen on the right-hand side of Figure 8. Note also how the number of balls intersecting decreases, when compared to the regular closed balls on the left-hand side. This leads to the dimension of the simplices in the Alpha complex to decrease as well. Moreover, we get that

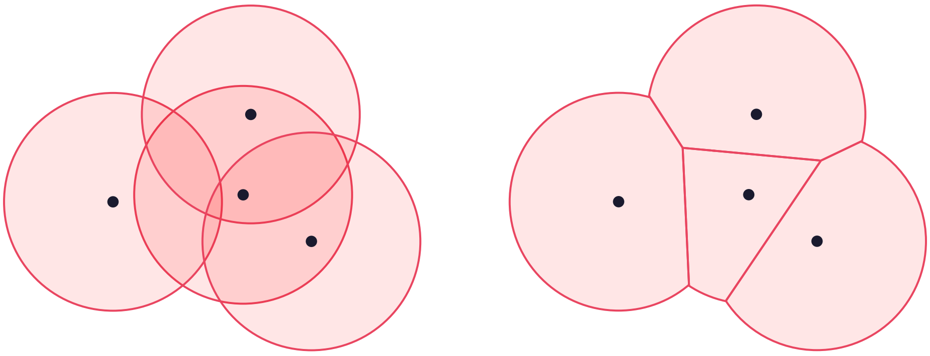

Figure 8: The regular closed balls are shown on the left and the Voronoi balls are shown on the right, given four points.

$\mathrm{Vor}_\infty(X, x) = \mathrm{Vor}(X, x)$, which means $\mathrm{Alpha}_\infty(X) = \mathrm{Del}(X)$. Hence, there is a filtration of $\mathrm{Del}(X)$ formed by Alpha complexes since $r \leq s$ implies $\mathrm{Alpha}_r(X) \subseteq \mathrm{Alpha}_s(X)$. This filtration gives us the following corollary of Proposition 2.10.

Corollary 2.13. *If at most $d + 1$ points in a finite set $X \subseteq \mathbb{R}^d$ lie on a common $(d-1)$-sphere, then the dimension of any simplex in* $\mathrm{Alpha}_r(X)$ *is at most d.*

One may find authors who call the Delaunay complex the Delaunay triangulation and the Alpha complex the Delaunay complex – notably [3]. I have, however, decided to stay consistent with the terminology used in [5].

2.2.4 Selective Alpha Complexes

It turns out that the Alpha and Čech complexes are extremal cases of a more general simplicial complex. The purpose of this complex will, first and foremost, be in proving the homotopy equivalences of some Euclidean data complexes in

Section 5. Let us start by generalizing the Voronoi cell of $x \in X$ to be defined with respect to an arbitrary subset E of X, i.e.,

$$\mathrm{Vor}(E, x) = \{p \in \mathbb{R}^d \mid d(x, p) \leq d(e, p) \text{ for all } e \in E\}.$$

Using this definition also gives us the *Voronoi ball with respect to* E, i.e.,

$$\mathrm{Vor}_r(E, x) = \mathrm{Vor}(E, x) \cap \bar{B}_d(x, r).$$

Definition 2.14. Given a finite set $X \subseteq \mathbb{R}^d$, then the *selective Alpha complex*[4] for $E \subseteq X$ and with radius $r \geq 0$ is the simplicial complex

$$\mathrm{Alpha}_r(X, E) = \left\{ \sigma \subseteq X \mid \bigcap_{x \in \sigma} \mathrm{Vor}_r(E, x) \neq \emptyset \right\}.$$

We write $\mathrm{Del}(X, E) = \mathrm{Alpha}_\infty(X, E)$ and call it the *selective Delaunay complex*. It is clear that if we are given $x \in X$, then $\mathrm{Vor}(\emptyset, x)$ will vacuously be all of $\mathbb{R}^d$, making $\mathrm{Vor}_r(\emptyset, x) = \bar{B}_d(x, r)$ for all $r \geq 0$. Then $\mathrm{Alpha}_r(X, \emptyset)$ is exactly the definition of $\check{\mathrm{C}}\mathrm{ech}_r(X)$. While if $E = X$, then $\mathrm{Alpha}_r(X, E) = \mathrm{Alpha}_r(X)$. The novelty must then come from the case when $\emptyset \subsetneq E \subsetneq X$. In this case, the Voronoi cells do not have to respect the boundaries of every other Voronoi cell. Only points in E respect the boundaries of each other's Voronoi cells. Hence, the result is a regular Voronoi diagram for points in E, with the remaining Voronoi cells overlaying it. See Figure 9 for an example. This, in turn, gives a simplicial

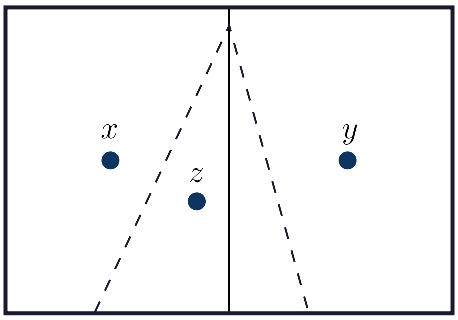

Figure 9: The Voronoi cell $\mathrm{Vor}(E, z)$ overlaps with $\mathrm{Vor}(E, x)$ and $\mathrm{Vor}(E, y)$ when $E = \{x, y\}$.

complex $\mathrm{Alpha}_r(X, E)$ that is somewhere in-between the Čech and Alpha complex. We can make this more precise.

[4]Note that due to the different choice in terminology, [3] introduced this as the *selective Delaunay complex*.

Lemma 2.15. *If $E \subseteq F$ are subsets of X, then $\mathrm{Vor}_r(F, x) \subseteq \mathrm{Vor}_r(E, x)$ for all $x \in X$.*

Proof. Recall that a point p lies in $\mathrm{Vor}(F, x)$ if and only if $d(p, x) \leq d(p, f)$ for every $f \in F$. In particular, $d(p, x) \leq d(p, e)$ for every $e \in E \subseteq F$. This gives us that $p \in \mathrm{Vor}(E, x)$. Then we get the desired result since $\mathrm{Vor}_r(G, x) \subseteq \mathrm{Vor}(G, x)$ for all $G \subseteq X$. □

As a consequence of Lemma 2.15, the Alpha and Čech complexes are extremal cases of the selective Alpha complex in the sense that

$$\underbrace{\mathrm{Alpha}_r(X, X)}_{\mathrm{Alpha}_r(X)} \subseteq \mathrm{Alpha}_r(X, E) \subseteq \underbrace{\mathrm{Alpha}_r(X, \emptyset)}_{\check{\mathrm{C}}\mathrm{ech}_r(X)}$$

for any $E \subseteq X$.

Bauer and Edelsbrunner [3] give another way to define the selective Alpha complex and to state Lemma 2.15, namely by means of *radius functions*. This formulation generalizes Delaunay spheres. In Section 4, we will that they are examples of *generalized discrete Morse functions*.

Definition 2.16. Given a $(d-1)$-sphere $S \subseteq \mathbb{R}^d$ and a finite set $X \subseteq \mathbb{R}^d$, let $\mathrm{Incl}\, S$ be the set of *included points* in X, i.e., the points that lie on or inside S. Similarly, let $\mathrm{Excl}\, S$ be the set of *excluded points* in X, i.e., the points that lie on or outside S.

We call the points in $\mathrm{On}\, S = \mathrm{Incl}\, S \cap \mathrm{Excl}\, S$ the points in X that are *on* S. We begin with subsets σ and E of a finite set $X \subseteq \mathbb{R}^d$. We will say that a $(d-1)$-sphere $S \subseteq \mathbb{R}^d$ *includes* σ if $\sigma \subseteq \mathrm{Incl}\, S$ and *excludes* E if $E \subseteq \mathrm{Excl}\, S$. If S includes σ and excludes E, while $A = \sigma \cap E$ is not empty, then clearly $A \subseteq \mathrm{On}\, S$. It is possible for there to be no spheres that both include σ and exclude E, but let $S(\sigma, E)$ be the sphere with the smallest radius r in the case there are such spheres. Define $s(\sigma, E)$ to be equal r^2. We call $S(\sigma, E)$ the *Delaunay sphere of σ with respect to E*. As expected, $S(\sigma, X)$ is the Delaunay sphere of σ we have already seen, while $S(\sigma, \emptyset)$ is the boundary of the miniball of σ. Hence, it makes sense to call the latter the *Čech sphere* of σ.

Definition 2.17. Let σ and E be subsets of the finite set $X \subseteq \mathbb{R}^d$. Then the *radius function*

$$s_E \colon \mathrm{Del}(X, E) \to \mathbb{R}$$

for E is defined by $s_E(\sigma) = s(\sigma, E)$.

We call s_X the *Delaunay radius function* of X and $s_\emptyset$ the *Čech radius function* of X. Just as the miniballs and Delaunay spheres were used to determine whether a simplex belongs to the Čech and Delaunay complexes, respectively, s_E will do the same for the selective Alpha complex. The proof is similar to Proposition 2.7.

Lemma 2.18. *Let $X \subseteq \mathbb{R}^d$ be a finite set and $E \subseteq X$. If given $r \geq 0$, then a simplex σ in $\mathrm{Del}(X, E)$ is also a simplex in $\mathrm{Alpha}_r(X, E)$ if and only if $s_E(\sigma) \leq r^2$.*

Proof. Suppose σ is a simplex in $\mathrm{Alpha}_r(X, E) \subseteq \mathrm{Del}(X, E)$ for some $r \geq 0$. By definition, the Voronoi balls with respect to E of the points in σ need to have some point y in common. For this to be the case, the furthest distance d between any point in $\sigma \cap E$ to y will necessarily be less than r. Moreover, the sphere of radius d with center y will include σ and exclude E. This implies that $s_E(\sigma) \leq d^2 \leq r^2$ since $s_E(\sigma)$ is a lower bound on d^2.

Conversely, suppose $s_E(\sigma) \leq r^2$ for some $r \geq 0$ and that σ is a simplex in $\mathrm{Del}(X, E)$. This implies that the center of the smallest sphere including σ and excluding E lies in

$$\bigcap_{x \in \sigma} \mathrm{Vor}_r(E, x).$$

Hence, we know that $\sigma \in \mathrm{Alpha}_r(X, E)$. $\square$

Using Lemma 2.18, we see that $s_E(\sigma) \leq s_F(\sigma)$ if and only if $\mathrm{Vor}_r(F, x) \subseteq \mathrm{Vor}_r(E, x)$. We can now restate Lemma 2.15 with the help of radius functions.

Corollary 2.19. *If $E \subseteq F$ are subsets of the finite set X, then $s_E(\sigma) \leq s_F(\sigma)$ whenever both functions are defined.*

2.2.5 Delaunay-Čech Complexes

The Alpha complex is defined by limiting the closed balls – used to define the Čech complex – to the Voronoi cells, which are used to define the Delaunay complex. The Delaunay-Čech complex, on the other hand, is defined by limiting the Čech *complex* itself to the Delaunay *complex*. As a result, both complexes are similar in that they use ideas from the Delaunay and Čech complexes to define a new complex.

Definition 2.20. Given a finite set $X \subseteq \mathbb{R}^d$, then the *selective Delaunay-Čech complex* for $E \subseteq X$ with radius $r \geq 0$ is the simplicial complex

$$\mathrm{Del\check{C}}_r(X, E) = \left\{ \sigma \subseteq \mathrm{Del}(X, E) \mid \bigcap_{x \in \sigma} \bar{B}_d(x, r) \neq \emptyset \right\}.$$

Note that we are mostly interested in the case when $E = X$, and call $\mathrm{Del\check{C}}_r(X) = \mathrm{Del\check{C}}_r(X, X)$ the *Delaunay-Čech complex*. It is only the latter complex that is considered in [3, 5], but making it selective will help motivate the next Euclidean data complex in this section.

But first we would like to compare the Alpha to the Delaunay-Čech complex. One way to distinguish them is to note that the latter only considers whether there is an intersection of closed balls at *any* point, as well as whether there is an intersection of the respective Voronoi cells at any other point. This does not have to be the same point. The Alpha complex, on the other hand, requires that the closed balls and the respective Voronoi cells do meet at the same point for a new simplex to be formed. Hence, in Figure 8 the Alpha complex would not contain a 3-simplex, while the Delaunay-Čech complex would (given that the Voronoi cells of the left and right points would eventually meet). One can think of this as saying the following. To form a simplex when we know some Voronoi cells meet, the Alpha complex requires the respective closed balls to have a radius r_0 that is sufficiently large for all the balls to reach a point at which the Voronoi cells meet. But for the Delaunay-Čech complex a simplex will be formed when the radius is large enough for the balls to meet at *some* point. It will at most require a radius of r_0. By this reasoning one can see that $\mathrm{Alpha}_r(X) \subseteq \mathrm{Del\check{C}}_r(X)$. Both complexes form a filtration of $\mathrm{Del}(X)$.

2.2.6 Selective Delaunay-Alpha Complexes

All the previous complexes are well-known Euclidean data complexes. We will now introduce a simplicial complex that we arrived at independently, but is implicit in [3, Theorem 5.9]. Its main purpose is to generalize all of the Euclidean data complexes we have defined in this thesis.

Definition 2.21. Given a finite set $X \subseteq \mathbb{R}^d$, then the *selective Delaunay-Alpha complex* for $E, F \subseteq X$ with radius $r \geq 0$ is the simplicial complex

$$\mathrm{DelA}_r(X, E, F) = \left\{ \sigma \subseteq \mathrm{Del}(X, E) \mid \bigcap_{x \in \sigma} \mathrm{Vor}_r(F, x) \neq \emptyset \right\}.$$

The new definition naturally leads us to ask whether the results proven in [3, 5] generalize to this complex as well. We will tackle this question in Section 6. For convenience, we gather all instances of the selective Delaunay-Alpha complex into Table 1. It illustrates how different choices of $E, F \subseteq X$ produce all of the Euclidean data complexes we have already defined. By making $r = \infty$ and $F = E$, we even get $\mathrm{Del}(X, E)$. Unfortunately, we will see that the selective Delaunay-Alpha dissimilarity will not be able to generalize the Delaunay dissimilarity since

r	E	F	$\mathrm{DelA}_r(X, E, F)$
r	$\emptyset$	$\emptyset$	$\check{\mathrm{C}}\mathrm{ech}_r(X)$
∞	E	E	$\mathrm{Del}(X, E)$
r	$\emptyset, F$	F	$\mathrm{Alpha}_r(X, F)$
r	E	$\emptyset$	$\mathrm{Del}\check{\mathrm{C}}_r(X, E)$

Table 1: Given a finite set $X \subseteq \mathbb{R}^d$, then choosing different subsets $E, F \subseteq X$ and radius $r \geq 0$ gives a different complex $\mathrm{DelA}_r(X, E, F)$.

we do not have the same control over r. This is, however, not a substantial problem.

As a last remark, we note that we clearly can define the selective Delaunay-Alpha complex with radius r as

$$\mathrm{DelA}_r(X, E, F) = \mathrm{Del}(X, E) \cap \mathrm{Alpha}_r(X, F).$$

This will help us show that this Euclidean data complexes is equivalent to all the others if $F \subseteq E$.

3 Dissimilarities

We begin by giving the general definition of a dissimilarity as well as some constructions associated with it. In the next section we turn to the specific dissimilarities we need when proving one of our main results.

3.1 Generalized Metrics

We generalize a metric on a topological space by removing its axioms and just considering continuous functions into $[0, \infty]$. Moreover, we want the ability to compare two different topological spaces X and Y, and say how dissimilar a point in X is from a point in Y.

Definition 3.1. Given topological spaces X and Y, a *dissimilarity* is a continuous function $\Lambda\colon X \times Y \to [0, \infty]$, where $[0, \infty]$ is given the order topology.

This is a noticeably a very general definition, but all our examples will have $Y = \mathbb{R}^d$ and X a finite subset of Y. Then we can think of X as our point cloud and define how dissimilar a point in $\mathbb{R}^d$ is from every point in X. We shall return to this when defining our Euclidean data complexes using dissimilarities.

Example 3.2. The most obvious example of a dissimilarity is a *metric* $D\colon X \times X \to [0, \infty)$ on a set X. If we let X be $\mathbb{R}^d$, then a metric that is particularly useful is the *Euclidean distance*

$$d(x, y) = \sqrt{\sum_{i=1}^{d} (y_i - x_i)^2}$$

between points $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ in $\mathbb{R}^d$. Note that we can find the distance between a point $p \in \mathbb{R}^d$ and a subset $X \subseteq \mathbb{R}^d$ if we let

$$d(p, X) = \inf\{d(p, x) \mid x \in X\}.$$

Definition 3.3. A *morphism* $f\colon \Lambda \to \Lambda'$ of dissimilarities $\Lambda\colon X \times Y \to [0, \infty]$ and $\Lambda'\colon X' \times Y' \to [0, \infty]$ is given by the pair (f_1, f_2) of continuous functions $f_1\colon X \to X'$ and $f_2\colon Y \to Y'$ such that

$$\Lambda'(f_1(x), f_2(y)) \leq \Lambda(x, y)$$

for all $(x, y) \in X \times Y$.

We can think of this as saying that two elements remain at least as similar after applying the morphism as before.

Example 3.4. In the case of metrics, a morphism of dissimilarities $d_1\colon X \times X \to [0, \infty)$ and $d_2\colon Y \times Y \to [0, \infty)$ corresponds to a *Lipschitz continuous* function $f\colon X \to Y$, where the *Lipschitz contsant* is 1. Note that an endomorphism of a single dissimilarity is also similar to the idea of a *contraction*, but where the "contraction constant" is 1 instead of strictly less than 1 [24, p. 220].

We can generalize the notion of a ball centered at some point to any dissimilarity.

Definition 3.5. Given a dissimilarity $\Lambda\colon X \times Y \to [0, \infty]$, then the Λ-*ball* of radius $r > 0$ centered at $x \in X$ is the subset

$$B_\Lambda(x, r) = \{y \in Y \mid \Lambda(x, y) < r\}$$

of Y.

Using the Euclidean metric d as our dissimilarity, we see that the definition of $B_d(x, r)$ given in the previous section coincides with the one above. We can think of this as thickening the point x to a radius r. If we do this for every point in X with respect to Y, then it can be considered a thickening of Λ.

Definition 3.6. The r-*thickening* of Λ is the subset

$$\Lambda^r = \bigcup_{x \in X} B_\Lambda(x, r)$$

of Y for $r > 0$.

3.2 The Dowker Nerve Theorem

Recall that we defined simplicial complexes by covering our point cloud with some balls and constructing an n-simplex where $n+1$ such balls intersect. This is called taking the *nerve* of a cover. One of the most central results in computational topology – and a big reason why the complexes in this thesis are interesting – is known as the *nerve theorem*. It tells us that the nerve of a convex covering is homotopic to the union of the sets in the cover [15]. In other words, we get all the combinatorial benefits of looking at simplicial complexes, while maintaining the homotopy type of the cover. We will state the nerve theorem in the context of dissimilarities along similar lines. To this end, we need to define the *Dowker nerve* of a dissimilarity, before stating the Dowker Nerve Theorem.

Definition 3.7. The *Dowker nerve* $N\Lambda$ of a dissimilarity $\Lambda\colon X \times Y \to [0,\infty]$ is the filtered simplicial complex $N\Lambda$ where

$$N\Lambda_r = \{\sigma \subseteq X \mid \text{there exists } y \in Y \text{ such that } \Lambda(x,y) < r \text{ for all } x \in \sigma\}$$

for $r > 0$.

Let $f\colon \Lambda \to \Lambda'$ be a morphism of dissimilarities $\Lambda\colon X \times Y \to [0,\infty]$ and $\Lambda'\colon X' \times Y' \to [0,\infty]$, given by (f_1, f_2), and let $\sigma \in N\Lambda_r$. If y is in Y such that $\Lambda(x,y) < r$ for every x in σ, then $\Lambda'(f_1(x), f_2(y)) < r$, by the definition of f. Hence, $f_1(\sigma)$ must be in $N\Lambda'$, which means f induces a simplicial map $Nf\colon N\Lambda \to N\Lambda'$.

Definition 3.8. A *partition of unity subordinate to the dissimilarity* $\Lambda\colon X \times Y \to [0,\infty]$ is a collection of continuous functions $\varphi^r\colon \Lambda^r \to |N\Lambda_r|$ such that

$$\overline{\{y \in Y \mid \varphi^r(y)(x) > 0\}} \subseteq B_\Lambda(x,r)$$

if x is in X.

Note that if Y is paracompact, then there always exists a partition of unity subordinate to $\Lambda\colon X \times Y \to [0,\infty]$ [12, p. 355]. Moreover, we say that a cover $\mathcal{U}$ of Y is *good* if every finite and non-empty intersection of sets in $\mathcal{U}$ is contractible.

Theorem 3.9 (The Dowker Nerve Theorem)**.** *Given a dissimilarity $\Lambda\colon X \times Y \to [0,\infty]$, where Y is paracompact, there exists a partition of unity*

$$\{\varphi^r\colon \Lambda^r \to |N\Lambda_r|\}$$

subordinate to Λ for some $r > 0$. If the cover $\{B_\Lambda(x,r)\}_{x\in X}$ of Λ^r is a good cover, then φ^r is a homotopy equivalence.

Proof. See Theorem 3 in [5]. □

It is particularly the last part that will become useful. Instead of proving that the geometric realizations of the Dowker nerve of two dissimilarities are homotopy equivalent, we can show it for the r-thickening of those dissimilarities. In other words, if we are given dissimilarities Λ and Λ', then Λ^r and $(\Lambda')^r$ are homotopy equivalent if and only if $|N\Lambda_r|$ and $|N\Lambda'_r|$ are – given the conditions in Theorem 3.9 are satisfied. But first we need to introduce the specific dissimilarities that we will need.

3.3 Euclidean Data Complexes from Dissimilarities

Now we come to the second way of defining our Euclidean data complexes. This definition of the Čech, Delaunay and Delaunay-Čech complexes was introduced by Morten Brun and Nello Blaser [6], and was inspired by the work of C. H. Dowker [13]. We give a definition of the selective Alpha complex along the same lines. In addition, we introduce a generalization of the Delaunay-Čech dissimilarity, which we call the *selective Delaunay-Alpha dissimilarity*. As we have seen, the Euclidean metric is an example of a dissimilarity. Thus, it is not hard to define the Čech dissimilarity in terms of it.

Definition 3.10. The *Čech dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ is defined as the function

$$\check{\mathrm{C}}\mathrm{ech}^X : X \times \mathbb{R}^d \to [0, \infty],$$

where $\check{\mathrm{C}}\mathrm{ech}^X(x, p)$ is the usual Euclidean distance $d(x, p)$ between x and p in $\mathbb{R}^d$.

The Čech dissimilarity serves as a good example of how the definitions in this section relate to the standard definitions of Euclidean data complexes. If given $r > 0$, then the Dowker nerve of the Čech dissimilarity contains a simplicial complex $N\check{\mathrm{C}}\mathrm{ech}_r^X$.

Proposition 3.11. *The simplicial complex* $N\check{\mathrm{C}}\mathrm{ech}_r^X$ *is equal to* $\check{\mathrm{C}}\mathrm{ech}_r(X)$.

Proof. A subset $\sigma \subseteq X$ is a simplex in $N\check{\mathrm{C}}\mathrm{ech}_r^X$ if and only if there exists a point $p \in \mathbb{R}^d$ such that $\check{\mathrm{C}}\mathrm{ech}^X(x, p) = d(x, p) < r$ for every point x in σ. The latter is just Proposition 2.7, finishing the proof.[5] □

Let us now turn to the different Delaunay complexes and how we define their dissimilarities. The definitions should make it clear that all we are doing is making every point in our point cloud X more or less dissimilar to every point in $\mathbb{R}^d$. In the case of Delaunay dissimilarities, that means whether $p \in \mathbb{R}^d$ is in the Voronoi cell of $x \in X$ or not. This is a binary relationship, unlike the Čech complex. Let $\mathbb{R}^d_D$ be Euclidean space with the discrete topology.

Definition 3.12. The *discrete Delaunay dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ is the function

$$\mathrm{del}^X : X \times \mathbb{R}^d_D \to [0, \infty]$$

[5] Note that we here have strict inequality. This does make the nerve of the Čech dissimilarity different from the Čech complex in some very specific cases, but since a tiny perturbation of r makes them equal again, we do not think this is a meaningful difference.

defined by

$$\mathrm{del}^X(x,p) = \begin{cases} 0 & \text{if } p \in \mathrm{Vor}(X,x) \\ \infty & \text{otherwise.} \end{cases}$$

It is not hard to see that $N\mathrm{del}_r^X$ is just a less direct, but nonetheless equivalent, definition of $\mathrm{Del}(X)$, where $r > 0$. However, del^X is not continuous with respect to the Euclidean topology on $\mathbb{R}^d$. This means that we cannot make use of the nerve theorem as we would like. We will, hence, construct a version of the Delaunay dissimilarity that will be continuous with respect to the Euclidean topology. To this end, we will need to "ramp up" continuously from 0 to ∞ in such a way that the Dowker nerve remains the same. We need to find some threshold ϵ that says how much we can *thicken* a Voronoi cell before constructing a simplex not in $\mathrm{Del}(X)$. We obtain ϵ as follows.

If $\sigma \subseteq X$ is not in $\mathrm{Del}(X)$, then

$$\epsilon_\sigma := \inf_{p \in \mathbb{R}^d} \max\{d(p, \mathrm{Vor}(X,x)) \mid x \in \sigma\}$$

is a strictly positive real. The only alternative is for it to be zero, but that would mean every $x \in \sigma$ lies in the same simplex in $\mathrm{Del}(X)$, contradicting the assumption that $\sigma \notin \mathrm{Del}(X)$. In fact, ϵ_σ may be thought of as the distance to the midpoint of the vertices in X. Hence, we may choose an ϵ such that $\epsilon < \epsilon_\sigma/2$ for every $\sigma \subseteq X$ that is not in $\mathrm{Del}(X)$.

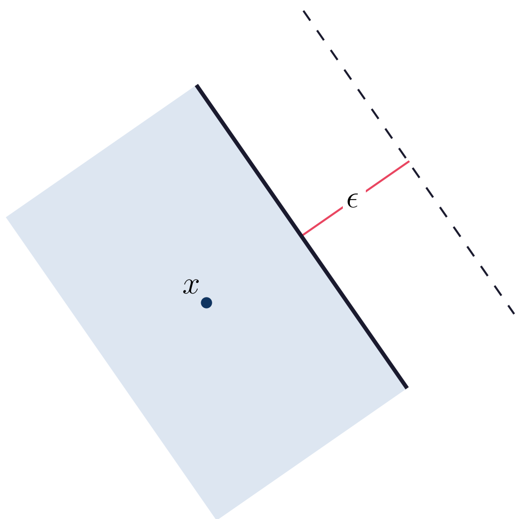

Figure 10: The ϵ-thickened Voronoi cell $\mathrm{Vor}(X,x)^\epsilon$.

Definition 3.13. Given $x \in X$ and an ϵ as found above, the ϵ-*thickened* Voronoi cell is the set

$$\mathrm{Vor}(X,x)^\epsilon = \{p \in \mathbb{R}^d \mid d(p, \mathrm{Vor}(X,x)) < \epsilon\}.$$

Figure 10 illustrates how we can think of $\mathrm{Vor}(X,x)^\epsilon$ geometrically. Now we have that the nerve of the open cover $\{\mathrm{Vor}(X,x)^\epsilon\}_{x\in X}$ of $\mathbb{R}^d$ is precisely $\mathrm{Del}(X)$. That is to say, we have thickened the Voronoi cells by ϵ to give ourselves a buffer to transition continuously from 0 to ∞, but with an ϵ small enough so that the nerve of $\{\mathrm{Vor}(X,x)^\epsilon\}_{x\in X}$ and $\{\mathrm{Vor}(X,x)\}_{x\in X}$ remain the same. The continuous transition is given by the following function.

Let $h\colon [0,\infty] \to [0,\infty]$ be the increasing function defined by

$$h(t) = \begin{cases} -\ln(1 - t/\epsilon) & \text{if } t < \epsilon \\ \infty & \text{if } t \geq \epsilon. \end{cases}$$

We have drawn a graph of h in Figure 11. Note how h serves as a "quick" way to go continuously from 0 to ∞.

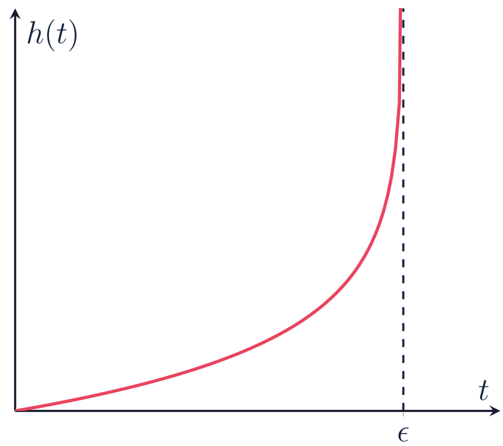

Figure 11: A graph of h, where $h(t) = \infty$ for $t \geq \epsilon$.

We describe the *generalized inverse* of h since that will be used in proving one of our main results. Given a subset $A \subseteq \mathbb{R}$, let $\overline{A}$ be its closure, and let $\overline{\mathbb{R}}$ denote the extended real numbers $\mathbb{R} \cup \{-\infty, \infty\}$.

Definition 3.14. Given an increasing function $f\colon A \to B$, where A and B are convex subsets of $\overline{\mathbb{R}}$, the *generalized inverse* $f^-\colon B \to \overline{A}$ of f is defined by

$$f^-(y) = \inf\{x \in A \mid f(x) \geq y\},$$

where we let $\inf \emptyset = \infty$.

Note that for a continuous and strictly increasing function we have that $f^{-} = f^{-1}$. Hence, the generalized inverse of h is given by

$$h^{-}(t) = \begin{cases} h^{-1}(t) & \text{if } t < \infty \\ \epsilon & \text{if } t = \infty. \end{cases}$$

We have drawn a graph of h^{-} in Figure 12. See [16] for more about generalized inverses.

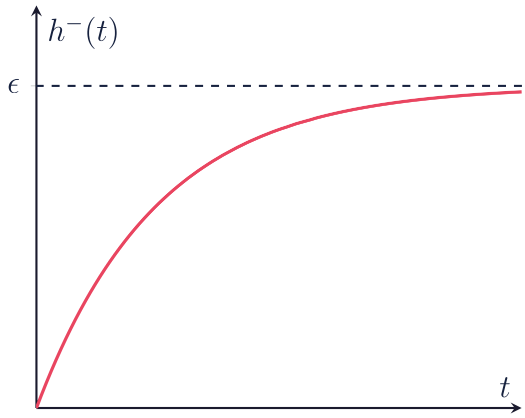

Figure 12: A graph of h^{-}, where $h^{-}(\infty) = \epsilon$.

Given $x \in X$, let $\mathrm{Del}_x \colon \mathbb{R}^d \to [0, \infty]$ be the function defined by

$$\mathrm{Del}_x(p) = h(d(p, \mathrm{Vor}(X, x))).$$

We see that $\mathrm{Del}_x(\mathrm{Vor}(X, x)) = 0$ and $\mathrm{Del}_x(\mathbb{R}^d - \mathrm{Vor}(X, x)^{\epsilon}) = \infty$. We can finally define the non-discrete version of the Delaunay dissimilarity.

Definition 3.15. The *Delaunay dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ is the function

$$\mathrm{Del}^X \colon X \times \mathbb{R}^d \to [0, \infty]$$

defined by

$$\mathrm{Del}^X(x, p) = \mathrm{Del}_x(p).$$

In order to simplify notation, we will write $N\mathrm{del}_r^X$ and $N\mathrm{Del}_r^X$ as respectively $N\mathrm{del}^X$ and $N\mathrm{Del}^X$ since their Dowker nerves remains unchanged for all $r > 0$.

Definition 3.16. The *discrete selective Delaunay dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ with respect to $E \subseteq X$ is the function

$$\mathrm{del}_E^X \colon X \times \mathbb{R}_D^d \to [0, \infty]$$

defined by

$$\operatorname{del}_E^X(x,p) = \begin{cases} 0 & \text{if } p \in \operatorname{Vor}(E,x) \\ \infty & \text{otherwise.} \end{cases}$$

As with the regular Delaunay dissimilarity, we need to construct a continuous version to use the Dowker Nerve Theorem 3.9. As before, choose a suitable ϵ for h, but this time depending on $N\operatorname{del}_E^X = \operatorname{Del}(X, E)$ instead of $\operatorname{Del}(X)$.

Definition 3.17. The *selective Delaunay dissimilarity* of X is the function

$$\operatorname{Del}_E^X \colon X \times \mathbb{R}^d \to [0, \infty]$$

defined by

$$\operatorname{Del}_E^X(x,p) = h(d(p, \operatorname{Vor}(E,x))).$$

By similar reasoning as before, we have that $N\operatorname{Del}_E^X = N\operatorname{del}_E^X$. Having defined both the Čech and Delaunay dissimilarities, we can also give the *Alpha dissimilarity.* However, to motivate the definition, we prove the following lemma first.

Lemma 3.18. *If* $\Lambda, \Lambda' \colon X \times Y \to [0, \infty]$ *are two dissimilarities, then*

$$\max(\Lambda, \Lambda') \colon X \times Y \to [0, \infty]$$

is also a dissimilarity, and

$$N \max(\Lambda, \Lambda')_r = N\Lambda_r \cap N\Lambda'_r$$

for $r > 0$*.*

Proof. The function $\max(\Lambda, \Lambda')$ is continuous since it is the maximum of two continuous functions, and hence a dissimilarity. For the second part, suppose σ is a simplex in $N \max(\Lambda, \Lambda')_r$, i.e., $\sigma \subseteq X$ such that there is some $y \in Y$ with $\max(\Lambda, \Lambda')(x, y) < r$ for all $x \in \sigma$. However, $\max(\Lambda, \Lambda')(x, y)$ is less than r if and only if $\Lambda(x, y) < r$ and $\Lambda'(x, y) < r$. This makes σ a simplex in $N\Lambda_r$ and $N\Lambda'_r$. Hence, we get the inclusion

$$N \max(\Lambda, \Lambda')_r \subseteq N\Lambda_r \cap N\Lambda'_r$$

of simplicial complexes. The other direction is equivalent. □

Definition 3.19. Given $E \subseteq X$, the *selective Alpha dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ is defined as

$$\operatorname{Alpha}_E^X = \max(\operatorname{Del}_E^X, \check{\operatorname{C}}\mathrm{ech}^X).$$

Using Lemma 3.18 and the definition above, we claim that the Dowker nerve of the selective Alpha dissimilarity

$$N(\mathrm{Alpha}_E^X)_r = N(\mathrm{Del}_E^X)_r \cap N\check{\mathrm{C}}\mathrm{ech}_r^X$$

gives us the selective Alpha complex $\mathrm{Alpha}_r(X)$. Since we have shown that the Dowker nerves of the Delaunay and Čech dissimilarities give us their corresponding complexes, we are left with showing that

$$\mathrm{Alpha}(X, E) = \mathrm{Del}_r(X, E) \cap \check{\mathrm{C}}\mathrm{ech}_r(X).$$

But this should be clear from the definition of $\mathrm{Alpha}_r(X, E)$. Being consistent with the corresponding simplicial complexes, we get $\mathrm{Alpha}_\emptyset^X = \check{\mathrm{C}}\mathrm{ech}^X$, and we define $\mathrm{Alpha}^X = \mathrm{Alpha}_X^X$ to be the *Alpha dissimilarity* of X.

In order to remain general, we will define a selective version of the Delaunay-Čech complex using dissimilarities. We will make use of both the discrete and continuous versions in proving one of our main results.

Definition 3.20. The *discrete selective Delaunay-Čech dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ with respect to $E \subseteq X$ is the function

$$\mathrm{del}\check{\mathrm{C}}_E^X \colon X \times (\mathbb{R}_D^d \times \mathbb{R}^d) \to [0, \infty]$$

defined by

$$\mathrm{del}\check{\mathrm{C}}_E^X(x, (p, q)) = \max\{\mathrm{del}_E^X(x, p), \check{\mathrm{C}}\mathrm{ech}^X(x, q)\}.$$

Note that we do not take the maximum of two functions, but of the two points $\mathrm{del}_E^X(x, p)$ and $\check{\mathrm{C}}\mathrm{ech}^X(x, q)$ in $\mathbb{R}$. Hence, taking the Dowker nerve of $\mathrm{del}\check{\mathrm{C}}_E^X$, this ensures that we are looking for an intersection of del_E^X-balls at some point, and $\check{\mathrm{C}}\mathrm{ech}^X$-balls at some – not necessarily the same – point. The definition is similar in the continuous case, except we use the continuous Delaunay dissimilarity.

Definition 3.21. The *selective Delaunay-Čech dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ with respect to $E \subseteq X$ is the function

$$\mathrm{Del}\check{\mathrm{C}}_E^X \colon X \times (\mathbb{R}^d \times \mathbb{R}^d) \to [0, \infty]$$

defined by

$$\mathrm{Del}\check{\mathrm{C}}_E^X(x, (p, q)) = \max\{\mathrm{Del}_E^X(x, p), \check{\mathrm{C}}\mathrm{ech}^X(x, q)\}.$$

We get the *(discrete) Delaunay-Čech dissimilarity* by letting $E = X$ in the definitions above, and denote them by respectively $\mathrm{del}\check{\mathrm{C}}^X$ and $\mathrm{Del}\check{\mathrm{C}}^X$. The simplicial complex $N\mathrm{del}\check{\mathrm{C}}_r^X = N\mathrm{Del}\check{\mathrm{C}}_r^X$ is equal to $\mathrm{Del}\check{\mathrm{C}}_r(X)$, by the argument above.

Continuing the idea of combining dissimilarities in order to create new ones, we arrive at the most general dissimilarity. Taking the Dowker nerve of this dissimilarity will give us a filtration of selective Delaunay-Alpha complexes.

Definition 3.22. The *selective Delaunay-Alpha dissimilarity* of a finite set $X \subseteq \mathbb{R}^d$ with respect to subsets $E, F \subseteq X$ is the function

$$\mathrm{DelA}^X_{E,F} \colon X \times (\mathbb{R}^d \times \mathbb{R}^d) \to [0, \infty]$$

defined by

$$\mathrm{DelA}^X_{E,F}(x, (p, q)) = \max\{\mathrm{Del}^X_E(x, p), \mathrm{Alpha}^X_F(x, q)\}.$$

Just as the selective Delaunay-Alpha complex generalizes many Euclidean data complexes, the selective Delaunay-Alpha dissimilarity generalizes most of the corresponding dissimilarities. Note that it cannot generalize the selective Delaunay dissimilarity, but the Dowker nerve does contain the complex

$$(N\mathrm{DelA}^X_{E,E})_\infty = \mathrm{Del}(X, E).$$

Analogous to what we did in Table 1, we collect the different instances of the selective Delaunay-Alpha dissimilarity.

Proposition 3.23. *Given a finite set $X \subseteq \mathbb{R}^d$ and subsets $E, F \subseteq X$, then*

$$\mathrm{DelA}^X_{E,F} = \begin{cases} \check{\mathrm{C}}\mathrm{ech}^X & \text{if } E = F = \emptyset \\ \mathrm{Del}\check{\mathrm{C}}^X_E & \text{if } F = \emptyset \\ \mathrm{Alpha}^X_F & \text{if } E = \emptyset, F. \end{cases} \tag{1}$$

This is not telling us anything new. Taking the Dowker nerve of Equation (1) gives us a filtration of the corresponding complexes in Table 1. It tells us that what we did in this section is just a different way to understand Euclidean data complexes. It allows us to show that some of the Euclidean data complexes are equivalent in Section 5 using tools such as r-thickened dissimilarities and the Dowker nerve theorem. In the next section, we will introduce another toolbox for proving that two simplicial complexes are equivalent. This will allow us to prove that *all* the Euclidean data complexes from Section 2 are equivalent.

4 Discrete Morse Theory

Most of the discussion about discrete Morse theory in this section is adopted from [18, 19] and the combinatorial description in [10]. Write $\sigma^{(p)}$ if σ is a p-simplex. We say that a face $\tau^{(p-1)} \subseteq \sigma^{(p)}$ is a *facet*. Note that discrete Morse theory is more generally developed for CW-complexes [19], but will only be described for simplicial complexes.

4.1 Discrete Morse Functions

Definition 4.1. Given a simplicial complex K, a function $f\colon K \to \mathbb{R}$ is a *discrete Morse function* on K if for every $\sigma^{(p)} \in K$ we have

(1) $|\{\tau^{(p+1)} \supsetneq \sigma \mid f(\tau) \leq f(\sigma)\}| \leq 1$ and

(2) $|\{\upsilon^{(p-1)} \subsetneq \sigma \mid f(\upsilon) \geq f(\sigma)\}| \leq 1$.

In other words, a function is a discrete Morse function on K if every simplex $\sigma \in K$ is the facet of at most one simplex with lower or equal value, and if there is at most one facet of σ that has a higher or equal value. This leads to a natural example of a discrete Morse function.

Example 4.2. The function in Figure 13 is a discrete Morse function, given as the height of the simplicial complex K. We see that $\tau_3 \supsetneq \sigma_2, \sigma_3$ while $f(\tau_3)$ is greater than $f(\sigma_2) = f(\sigma_3)$, and $\sigma_1 \subsetneq \tau_1, \tau_2$ while $f(\sigma_1)$ less than $f(\tau_1) = f(\tau_2)$. Hence, the map f in Figure 13 does satisfy the conditions for it to be a discrete Morse function, as claimed.

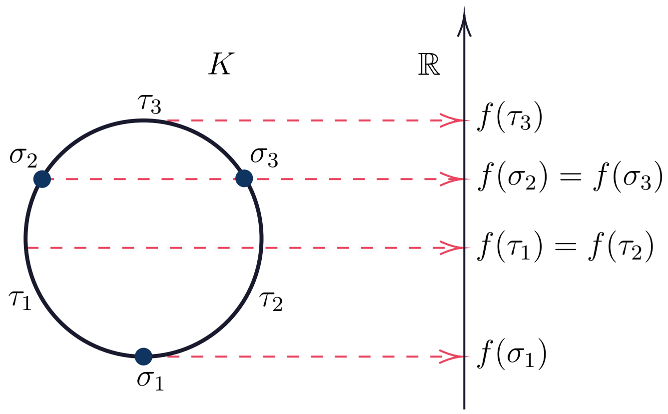

Figure 13: A discrete Morse function f given as a height function on K.

Definition 4.3. Given a discrete Morse function $f\colon K \to \mathbb{R}$, a simplex $\sigma^{(p)} \in K$ is *critical* with *critical value* $f(\sigma)$ if

(1) $|\{\tau^{(p+1)} \supsetneq \sigma \mid f(\tau) \leq f(\sigma)\}| = 0$ and

(2) $|\{\upsilon^{(p-1)} \subsetneq \sigma \mid f(\upsilon) \geq f(\sigma)\}| = 0$.

The function in Example 4.2 has σ_1 and τ_3 as critical simplices. The vertex σ_1 satisfies (1) since it is not the facet of a simplex with a *lower* value than σ_1, while (2) is vacuously true. Similarly, the edge τ_3 satisfies (2) since no facet of τ_3 has a *higher* value, and (1) is vacuously true. Using the definition of a Morse function, we get that any simplex must satisfy at least one non-vacuous test for criticality. This will play a crucial role when we talk about gradient vector fields of discrete Morse functions.

Lemma 4.4. *If K is a simplicial complex with a Morse function f, then for any simplex $\sigma^{(p)} \in K$ we must have*

(1) $|\{\tau^{(p+1)} \supsetneq \sigma \mid f(\tau) \leq f(\sigma)\}| = 0$, *or*

(2) $|\{\upsilon^{(p-1)} \subsetneq \sigma \mid f(\upsilon) \geq f(\sigma)\}| = 0$.

Proof. See Lemma 2.5 in [19]. □

Similarly to what is done in classical Morse theory, we can construct a simplicial complex by building it up with subcomplexes in the order given by the discrete Morse function. In other words, we define a filtration on a simplicial complex that has been given a discrete Morse function.

Definition 4.5. Given a simplicial complex K and a discrete Morse function f on K, the *level subcomplex* is

$$K(t) := \bigcup_{f(\sigma) \leq t} \bigcup_{\tau \subseteq \sigma} \tau$$

for $t \in \mathbb{R}$.

Note that since we are working with finite simplicial complexes, there are either no, or a finite number of values $f(\sigma) \leq t$ for a given $t \in \mathbb{R}$. So if we choose $t, s \in \mathbb{R}$ such that

$$t < \min_{\sigma \in K} f(\sigma) \quad \text{and} \quad s = \max_{\sigma \in K} f(\sigma),$$

then $K(t) = \emptyset$ and $K(s) = K$. Hence, we can define a finite increasing sequence $\{t_i\}_{i=0}^{n}$ from all values $f(\sigma)$ with $\sigma \in K$ and the t chosen above. This gives a filtration of K by letting $K_i := K(t_i)$ for $i = 0, 1, \ldots, n$.

A proper face τ of a maximal simplex $\sigma \in K$ is said to be *free* if τ is not the face of any other simplex in K. If L is a subcomplex of K such that $L = K - \{\sigma, \tau\}$, where τ is a free facet of σ in K, then we say K *collapses onto* L, and write $K \searrow L$ for the *simplicial collapse* itself. In general, we write $K \searrow L$ if there is a finite sequence

$$K = M_1 \searrow M_2 \searrow \cdots \searrow M_n = L$$

of such collapses. See Figure 14 for the collapse of a 2-simplex onto a vertex.

It is clear that – after taking the geometric realization – a simplicial collapse is an example of a deformation retract. In fact, such collapses were studied by J.H.C. Whitehead in the general case of CW-complexes. He called a homotopy equivalence generated by simplicial collapses for a *simple-homotopy equivalence*. Hence, one of the proofs in Section 5 will technically be stronger because of this. We refer the reader to [11] for more information about simple-homotopy theory. For our purposes, it will be sufficient to consider these as regular homotopies.

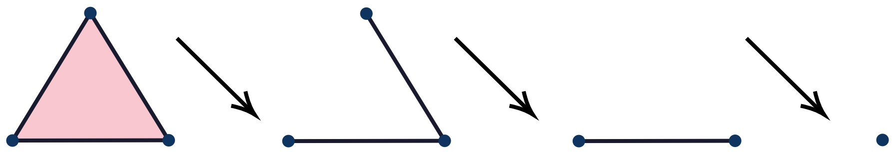

Figure 14: The simplicial collapse of a 2-simplex to a vertex.

Theorem 4.6 (Collapsing Theorem)**.** *Let $f\colon K \to \mathbb{R}$ be a discrete Morse function and $(a, b] \subseteq \mathbb{R}$. If there are no critical simplices $\sigma \in K$ with $f(\sigma) \in (a, b]$, then $K(b) \searrow K(a)$.*

Proof. See [19, Theorem 3.3]. □

Looking at Example 4.2, Theorem 4.6 says that we may collapse $K(f(\sigma_2))$ onto $K(f(\sigma_1))$. That is the same as saying that one may contract an arc to a point without changing its homotopy type since $(f(\sigma_1), f(\sigma_2)]$ does not contain a critical value. However, we cannot collapse $K(f(\tau_3))$ onto $K(f(\sigma_1))$ since $(f(\sigma_1), f(\tau_3)]$ now *does* contain a critical value, namely $f(\tau_3)$. This tells us, as expected, that a circle and a point are not (simple-)homotopy equivalent.

4.2 Discrete Vector Fields

Defining a discrete Morse function $f\colon K \to \mathbb{R}$ for nontrivial simplicial complexes can be very cumbersome. To avoid having to do this explicitly, we can rather look at its *discrete gradient* V, which can be thought of as the discrete analogue of the negative gradient vector field $-\nabla f$ in the smooth case. We begin more generally with an arbitrary *discrete vector field*.

Definition 4.7. A *discrete vector field* V on K is a subset of $K \times K$ such that each simplex of K is in at most one pair in V, and $(\sigma, \tau) \in V$ implies σ is a facet of τ.

We can think of V as containing arrows (σ, τ) with a tail σ and head τ. Then this tells us that a discrete vector field makes every simplex at most either the tail of a cofacet or the head of a facet. The discrete vector fields we are interested in, however, need to somehow depend on f. Since we are looking for the discrete analogue of $-\nabla f$, it is natural to also require for a discrete gradient V to define "vectors" (σ, τ) that go from a higher to a lower Morse value, i.e., we require that $f(\tau) \leq f(\sigma)$. In fact, there is a unique such V, hence justifying the following definition.

Proposition 4.8. *Let $f\colon K \to \mathbb{R}$ be a discrete Morse function. There exists a unique discrete vector field V such that $(\sigma, \tau) \in V$ if and only if σ is a facet of τ and $f(\tau) \leq f(\sigma)$.*

Proof. We only need to check that every simplex of K is in at most one pair in V for V to be a discrete vector field. Without loss of generality, let us assume σ is contained in the pair (σ, τ) in V. First suppose it is also in the pair (σ, υ) in V. Then, by definition, σ is a facet of both τ and υ with $f(\tau), f(\upsilon) \leq f(\sigma)$. However, condition (1) in the definition of a discrete Morse function states that there can be at most one simplex with such a property. Hence, we must conclude that $\tau = \upsilon$.

Now suppose σ is contained in the pair (υ, σ) in V. Then, as before, we know that υ is a facet of σ, which again is a facet of τ. Moreover, we get that $f(\tau) \leq f(\sigma) \leq f(\upsilon)$. This leads to a contradiction since Lemma 4.4 says that only one of the inequalities can be true.

To prove uniqueness, let us suppose there exists another discrete vector field V' such that (σ, τ) lies in V' if and only if σ is a facet of τ and $f(\tau) \leq f(\sigma)$. But the latter part is true if and only if (σ, τ) lies in V. The other containment is similar. $\square$

We call this discrete vector field the *discrete gradient* of f. Note that defining the discrete gradient in such a way also gives us the desired property that a critical simplex σ is not the head nor tail of any vector in V. In the smooth case this is usually stated as $\nabla f(\sigma) = 0$. The idea is that an arrow exists between two simplices precisely when there is a corresponding simplicial collapse. Hence, Theorem 4.6 says that the discrete gradient encodes all the simplicial collapses. This will become formalized when we reformulate the theorem.

Example 4.9. Since $-\nabla f$ defines vectors pointing in the direction of steepest descent, we might expect the same to be true of V. Looking at Example 4.2 once again, we have drawn its discrete gradient $V = \{(\sigma_2, \tau_1), (\sigma_3, \tau_2)\}$ in Figure 15.

Note how τ_3 in some way acts as a source and σ_1 acts as a sink, even if neither is the head nor tail of any arrow. This is consistent with the idea that the vectors in V point in the direction of steepest descent. Moreover, the two arrows correspond precisely to the simplicial collapse that we considered at the end of Section 4.1.

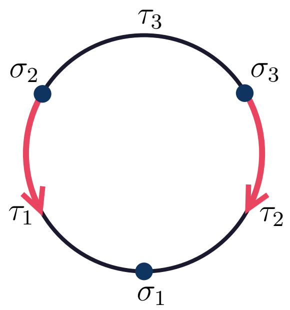

Figure 15: The discrete gradient V.

As hinted at in the beginning of the section, there is a neat combinatorial way of looking at the discrete vector fields using Hasse diagrams that is due to Chari [10]. Then we can look at discrete Morse functions that have a particular discrete vector fields as its discrete gradient. We begin by specifying discrete vector fields based on the cofacet relation. We organize this information into a single object.

Definition 4.10. The *Hasse diagram* $\mathcal{H}(K)$ of a simplicial complex K is the transitive reduction of the partially ordered set of simplices defined by the coface relation.

In other words, the coface relation defines a partially ordered set on K, which gives us $\mathcal{H}(K)$ when we remove all coface relations that are the composition of any other two. This means that $\mathcal{H}(K)$ only contains (τ, σ) if σ is a facet of τ. Clearly, this defines a directed acyclic graph $G(K)$ with K as the vertex set and $\mathcal{H}(K)$ the set of edges. For our reoccurring example, we get the graph in Figure 16.

$$\begin{tikzcd}
\tau_1 \arrow[d] \arrow[dr] & \tau_2 \arrow[dl] \arrow[dr] & \tau_3 \arrow[dl] \arrow[d] \\
\sigma_1 & \sigma_2 & \sigma_3
\end{tikzcd}$$

Figure 16: The directed acyclic graph $G(K)$ representing $\mathcal{H}(K)$. The dimension of the simplices increases upwards.

For the remainder of this thesis, it will be more convenient to work with discrete gradients as a partition of a simplicial complex K rather than a set of ordered pairs. Let a *discrete vector field* V be a partition of K into singleton sets $\{\sigma\}$ and unordered pairs $\{\tau, \upsilon\}$ corresponding to some edges (τ, υ) in $G(K)$. We let $G_V(K)$ be the graph obtained by reversing the edges in $G(K)$ corresponding to the pairs in V. Suppose a function $f\colon K \to \mathbb{R}$ has values $f(\sigma) \leq f(\tau)$ if σ is a face of τ, and $f(\sigma) = f(\tau)$ if and only if $\{\sigma, \tau\}$ is a pair in V. Then f is a discrete Morse function – as defined in Definition 4.1 – with *discrete gradient* V. One can check that a discrete vector field V of a simplicial complex K is a discrete gradient if and only if $G_V(K)$ is acyclic [10].

We see that there are uncountably many discrete Morse functions with V as their discrete gradient by scaling values of f by some positive real. But every discrete Morse function only has one discrete gradient. Hence, a discrete gradient acts as an equivalence class of discrete Morse functions. So although requiring equal values for pairs in V makes us consider a smaller collection of discrete Morse functions, we are in fact just picking convenient representatives of said equivalence classes.

By design, a face can never have a higher value than its coface. So to check if a simplex σ is critical is equivalent to checking whether it has a facet τ, or a cofacet υ, with equal value under f. However, we have also required that to be the case if and only if either $\{\tau, \sigma\}$, or $\{\sigma, \upsilon\}$, is a pair in V. Hence, a simplex is critical if it does not belong to any pair in V. Equivalently, a simplex is critical if it belongs to a singleton set in V. Note that although the critical simplices remain the same, the critical values may differ based on the function f.

Example 4.11. We may reconstruct the discrete gradient V of K in Figure 15 based on the Hasse diagram $\mathcal{H}(K)$ in Figure 16. We choose the pairs in V to be $\{\sigma_2, \tau_1\}$ and $\{\sigma_3, \tau_2\}$, corresponding to the pairs (τ_1, σ_2) and (τ_2, σ_3) in $\mathcal{H}(K)$. Then we let the last two simplices be singleton sets. This gives us the discrete gradient

$$V = \{\{\sigma_2, \tau_1\}, \{\sigma_3, \tau_2\}, \{\sigma_1\}, \{\tau_3\}\}.$$

We draw the resulting graph $G_V(K)$ in Figure 17. The only difference from how

$$\begin{tikzcd}
\tau_1 \arrow[d] & \tau_2 \arrow[dl] & \tau_3 \arrow[d] \arrow[dl] \\
\sigma_1 & \sigma_2 \arrow[ul] & \sigma_3 \arrow[ul]
\end{tikzcd}$$

Figure 17: The graph $G_V(K)$ corresponding to the discrete gradient V. The simplex σ_1 is a sink, while τ_3 is a source.

we previously described V is the addition of the singleton sets and not making the pairs ordered. Defining a discrete Morse function f that has V as its discrete gradient can be done using the same values as before, except we now require the equalities $f(\sigma_2) = f(\tau_1)$ and $f(\sigma_3) = f(\tau_2)$. The critical simplices are those that do not belong to a pair in V, i.e., the singleton sets $\{\sigma_1\}$ and $\{\tau_3\}$, as expected.

A sink always exists, and it must be a critical 0-simplex. A source must also exist, but it does not have to be of a particular dimension. It is critical if and only if it is a maximal simplex. This fact will be useful in proving the gradient version of Theorem 4.6.

Theorem 4.12 (Gradient Collapsing Theorem)**.** *Let V be the discrete gradient of a simplicial complex K, and let L be a subcomplex of K. If $K - L$ is a union of pairs in V, then $K \searrow L$.*

Proof. Half of the simplices in $K - L$ must be source nodes in $G_V(K)$, having an edge to the other half. Choosing a sufficiently large p, suppose $\sigma^{(p)} \in K - L$ is a source simplex with an edge to a simplex $\tau^{(p+1)} \in K - L$ that is maximal in K. Such a p and τ exist because L is a simplicial complex. Note that a maximal simplex which is a source in $G_V(K)$ has to be a critical simplex. Thus, σ is not maximal and τ is not a source since they make up a pair in V. Then, by definition of a discrete gradient, σ is free since it is the facet of exactly one simplex τ, and that simplex is maximal. This gives us the simplicial collapse $K \searrow K - \{\sigma, \tau\}$. Proceeding inductively, from the higher to the lower dimensional simplices in $K - L$, we get

$$K \searrow K - (K - L) = L,$$

as desired. □

Thus, a discrete gradient does encode simplicial collapses. Theorem 4.12 makes no reference to a discrete Morse function, which means we no longer need to find the specific function that will give us a desired simplicial collapse. This is a major benefit of working with discrete gradients over the discrete Morse functions themselves.

4.3 Generalized Discrete Morse Theory

In order to generalize our notions in discrete Morse theory, we consider intervals of faces of a simplicial complex K.

Definition 4.13. Given two simplices $\sigma, \tau \in K$, an *interval* $[\sigma, \tau]$ is a subset of K given by $\{v \in K \mid \sigma \subseteq v \subseteq \tau\}$.

If $[\sigma, \tau]$ is nonempty, we call σ the *lower bound* and τ the *upper bound*. The following terminology is due to Bauer and Edelsbrunner [3], inspired by the work of Ragnar Freji [20].

Definition 4.14. A *generalized discrete vector field* W is a partition of K into intervals.

As one would hope, this definition has as a special case the partition of K into intervals of the form $[\sigma, \sigma] = \{\sigma\}$ and $[\sigma, \tau] = \{\sigma, \tau\}$, giving us the second definition of a discrete vector field from Section 4.2. However, no generalization of discrete Morse theory is complete without also generalizing the discrete Morse function. This can be done by constructing a function $f\colon K \to \mathbb{R}$ such that some generalized discrete vector field W will become the *generalized discrete gradient* of f.

Definition 4.15. If given a simplicial complex K and a generalized discrete vector field W, then a function $f\colon K \to \mathbb{R}$ is a *generalized discrete Morse function* on K with *generalized discrete gradient* W if

(1) σ being a face of τ implies $f(\sigma) \leq f(\tau)$, and

(2) $f(\sigma) = f(\tau)$ if and only if there exists an interval $[\upsilon_1, \upsilon_2] \in W$ such that $\sigma, \tau \in [\upsilon_1, \upsilon_2]$.

Clearly, a discrete Morse function is a generalized discrete Morse function where its generalized discrete gradient only consists of singletons and pairs. In this way, we have generalized both the discrete gradient and the discrete Morse function.

An interval in a generalized discrete gradient W associated to the generalized discrete Morse function $f\colon K \to \mathbb{R}$ is called *singular* if it only contains a single simplex $\sigma \in K$. Then we call σ a *critical simplex* and $f(\sigma)$ a *critical value* of K. This is consistent with how simplices in singleton sets of a discrete gradient are critical. The link between the generalized and non-generalized case can be made even more clear.

Given a generalized discrete gradient W, we can construct a discrete gradient V. The only intervals a non-generalized discrete gradient can contain are singletons and pairs. So we will let V refine every non-singular interval $[\sigma, \tau] \in W$ into pairs by choosing an arbitrary vertex $x \in \tau - \sigma$ and partitioning $[\sigma, \tau]$ into pairs of the form $\{\upsilon - \{x\}, \upsilon \cup \{x\}\}$ for $\upsilon \in [\sigma, \tau]$. We call V a *vertex refinement* of W. A simplex in K is critical if it does not belong to any pair in V. This means that W and V contain the same critical simplices and allows us to state Theorem 4.12 in the generalized case.

Theorem 4.16 (Generalized Gradient Collapsing Theorem)**.** *Let K be a simplicial complex with a generalized discrete gradient W, and let L be a subcomplex of K. If $K - L$ is a union of non-singular intervals in W, then $K \searrow L$.*

Proof. Suppose $K - L$ is a union of non-singular intervals in W. Construct a vertex refinement V of W as above. Then $K - L$ only contains pairs in V since it is comprised of non-singular intervals. This implies that $K \searrow L$ by Theorem 4.12. □

4.3.1 Radius Functions

The reason for developing all the theory in this section comes down to the radius functions we first met in Section 2.2.4. These are, namely, examples of generalized discrete Morse functions. Although we will only give an outline of the argument given in [3], it is a crucial step in showing that one can collapse some Euclidean data complexes onto others.

It turns out that we can use convex optimization to approach this problem. We begin by describing a Delaunay sphere $S(\sigma, E)$ with respect to E as the sphere with center $p \in \mathbb{R}^d$ and radius $r \geq 0$ that solves the following constrained minimization problem:

$$\begin{aligned} \min_{p,\,r} \quad & r^2 \\ \text{s.t.} \quad & d(p,x)^2 \leq r^2 \text{ for all } x \in \sigma, \\ & d(p,e)^2 \geq r^2 \text{ for all } e \in E. \end{aligned}$$

Recall that $S(\sigma, E)$ is defined as the smallest enclosing sphere including σ and excluding E. In other words, we want to find a center p and radius r that defines a sphere that has the smallest squared radius possible, subject to the conditions of inclusion and exclusion. This also defines the radius function s_E. If we now use Lemma 2.18, we can determine if a simplex σ belongs to $\text{Alpha}_r(X, E)$ via the convex optimization problem above.

We want to rephrase this problem to prove that the radius functions are generalized discrete Morse functions. In the field of constrained optimization, there are first derivative tests called the *Karush-Kuhn-Tucker (KKT) conditions* that are helpful. It is beyond the scope of this thesis to introduce these conditions in their full generality, but more details are provided in [3]. What is more relevant for us, is to see how we can use a specific example of these conditions to describe generalized discrete Morse functions.

We need to restrict ourselves to the case in which our point cloud $X \subseteq \mathbb{R}^d$ is in general position. It should be noted that this does make the final result less general. Recall that X being in general position tells us that the center p of the smallest circumscribed sphere S of $\sigma \subseteq X$ must lie in the affine hull of $\sigma = \operatorname{On} S$. In other words,

$$p = \sum_{x \in \operatorname{On} S} \rho_x x \quad \text{such that} \quad \sum_{x \in \operatorname{On} S} \rho_x = 1$$

and ρ_x is non-zero and unique for each $x \in \operatorname{On} S$.

Definition 4.17. Given a $(d-1)$-sphere $S \subseteq \mathbb{R}^d$ and a finite set $X \subseteq \mathbb{R}^d$, the *front face* of $\operatorname{On} S \subseteq X$ is the subset

$$\operatorname{Front} S = \{x \in \operatorname{On} S \mid \rho_x > 0\},$$

while the *back face* of $\operatorname{On} S$ is the subset

$$\operatorname{Back} S = \{x \in \operatorname{On} S \mid \rho_x < 0\}.$$

As an example, we draw the same circumscribed 1-sphere of three points in two different positions in Figure 18. Note that $\operatorname{Back} S$ is empty if the center p of

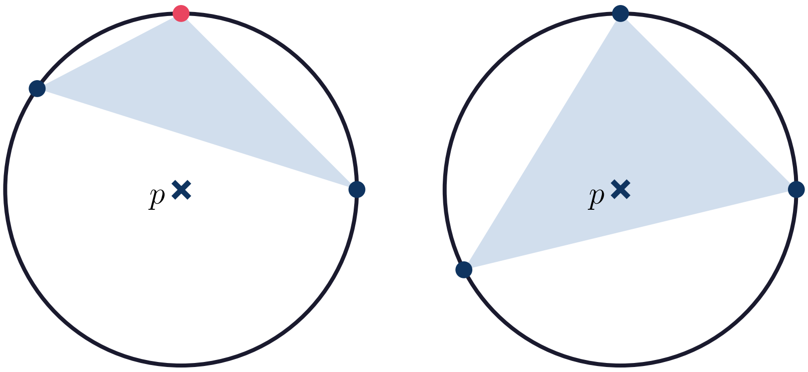

Figure 18: On the left, $\operatorname{Back} S$ contains the top point (shown in red), and $\operatorname{Front} S$ contains the other two in $\operatorname{On} S$. On the right, $\operatorname{Front} S$ contains all three points in $\operatorname{On} S$ since p is in the convex hull of $\operatorname{On} S$ (shaded blue).

S is in the convex hull of the points $\operatorname{On} S$ since, by definition, $\rho_x > 0$ for every $x \in \operatorname{On} S$. We can use $\operatorname{Front} S$ and $\operatorname{Back} S$ to state a combinatorial version of the KKT conditions.

Theorem 4.18 (Combinatorial KKT Conditions)**.** *Let $X \subseteq \mathbb{R}^d$ be a finite set in general position and let $\sigma, E \subseteq X$ be two subsets such that there exists a sphere $S \subseteq \mathbb{R}^d$ with $\sigma \subseteq \operatorname{Incl} S$ and $E \subseteq \operatorname{Excl} S$. Then $S = S(\sigma, E)$ if and only if*

(1) S is the smallest circumscribed sphere of $\operatorname{On} S$,

(2) $\operatorname{Front} S \subseteq \sigma$ and

(3) $\operatorname{Back} S \subseteq E$.

Proof. See Theorem 4.3 in [3]. □

Given a point cloud X and $\sigma, E \subseteq X$, Theorem 4.18 tells us when a sphere S is the Delaunay sphere of σ with respect to E. Hence, it can be used in conjunction with Lemma 2.18 to determine whether σ is a simplex in $\mathrm{Alpha}_r(X, E)$ for $r \geq 0$. This gives us all the tools necessary to prove that the radius functions are generalized discrete Morse functions.

Theorem 4.19 (Selective Delaunay Morse Function Theorem)**.** *Let $X \subseteq \mathbb{R}^d$ be a finite set in general position with $E \subseteq X$. Then the radius function*

$$s_E \colon \mathrm{Del}(X, E) \to \mathbb{R}$$

is a generalized discrete Morse function with the generalized discrete gradient

$$W = \{[\operatorname{Front} S(\sigma, E), \operatorname{Incl} S(\sigma, E)] \mid \sigma \in \mathrm{Del}(X, E)\}\,.$$

Proof. If $\tau \subseteq \sigma$ are simplices in $\mathrm{Del}(X, E)$ contained in two different intervals in W, then $s_E(\tau) < s_E(\sigma)$ since X is in general position, which means τ is not included by $S(\sigma, E)$. This is Condition (1) of Definition 4.15. Having fixed $E \subseteq X$ and $S = S(\sigma, E)$, note that $s_E(\tau) = s_E(\sigma)$ for all $\tau \in [\operatorname{Front} S, \operatorname{Incl} S]$. This ensures Condition (2), i.e., s_E only has equal values for simplices in the same interval in W. □

Just by changing $E \subseteq X$ in Theorem 4.19, we get that the Čech and Delaunay radius functions are also generalized discrete Morse. Recall that $\mathrm{Del}(X, \emptyset)$ is the full simplex $\Delta(X)$ and that $\mathrm{Del}(X, X) = \mathrm{Del}(X)$.

Corollary 4.20 (Čech Morse Function Corollary)**.** *Let $X \subseteq \mathbb{R}^d$ be a finite set in general position. The Čech radius function is a generalized discrete Morse function with generalized discrete gradient*

$$\{[\operatorname{On} S(\sigma, \emptyset), \operatorname{Incl} S(\sigma, \emptyset)] \mid \sigma \in \Delta(X)\}\,.$$

Proof. If we let $E = \emptyset$, then Theorem 4.18 forces $\operatorname{Back} S$ to be empty. Hence, we get $\operatorname{Front} S = \operatorname{On} S$. Using Theorem 4.19 gives us that $s_\emptyset$ is a generalized discrete Morse function with discrete gradient containing intervals of the form $[\operatorname{On} S, \operatorname{Incl} S]$ for all Čech spheres of X. □

Corollary 4.21 (Delaunay Morse Function Corollary)**.** *Let $X \subseteq \mathbb{R}^d$ be a finite set in general position. The Delaunay radius function is a generalized discrete Morse function with generalized discrete gradient*

$$\{[\operatorname{Front} S(\sigma, X), \operatorname{On} S(\sigma, X)] \mid \sigma \in \mathrm{Del}(X)\}\,.$$

Proof. This proof is similar. If we let $E = X$, then Theorem 4.18 forces $\operatorname{Front} S$ to be empty. Hence, we get $\operatorname{Incl} S = \operatorname{On} S$. Using Theorem 4.19 gives us that s_X is a generalized discrete Morse function with discrete gradient containing intervals of the form $[\operatorname{Front} S, \operatorname{On} S]$ for all Delaunay spheres of X. $\square$

This concludes most of the theory necessary to tackle the main results of this thesis. It will hopefully become clear in the following section why we discussed dissimilarities and discrete Morse theory so extensively. All the theory comes together to show – in two different ways – that all the Euclidean data complexes in Section 2 are equivalent.

5 Equivalent Euclidean Data Complexes

Thus far, we have dealt with a lot of definitions and concepts related to simplicial complexes. This section will bring these together into some important results, and hence deals with the main theorem of this thesis: the Alpha, Delaunay-Čech and Čech complexes are homotopy equivalent. Let us start by explicitly stating the theorem, and then discuss the two different approaches in proving it.

Theorem 5.1 (Equivalent Euclidean Data Complexes Theorem)**.** *Let $X \subseteq \mathbb{R}^d$ be a finite set. The Euclidean data complexes*

$$\mathrm{Alpha}_r(X) \subseteq \mathrm{Del\check{C}}_r(X) \subseteq \check{\mathrm{C}}\mathrm{ech}_r(X)$$

are homotopy equivalent[6] *for a given $r \geq 0$.*

The first proof of Theorem 5.1 is done directly using elementary geometry, while the second relies on generalized discrete Morse theory. The geometric argument [5] uses less theory than the gradient collapse argument [3] which gives a slightly weaker conclusion, but it does not assume X is in general position in return.

5.1 Geometric Argument

We begin with the proof by Nello Blaser and Morten Brun [5], as it is more intuitive and requires less theory, given we are already familiar with the definition of Euclidean data complexes in terms of dissimilarities. Recall that taking the Dowker nerve of the Alpha, Delaunay-Čech and Čech dissimilarities gives us their corresponding Euclidean data complexes. In particular, if we also take their geometric realization, we have $|N\mathrm{Alpha}_r^X| = |\mathrm{Alpha}_r(X)|$, $|N\check{\mathrm{C}}\mathrm{ech}_r^X| = |\check{\mathrm{C}}\mathrm{ech}_r(X)|$ and $|N\mathrm{Del\check{C}}_r^X| = |\mathrm{Del\check{C}}_r(X)|$. We will thus prove Theorem 5.1 if we can provide homotopy equivalences $(\mathrm{Alpha}^X)^r \to (\check{\mathrm{C}}\mathrm{ech}^X)^r$ and $(\check{\mathrm{C}}\mathrm{ech}^X)^r \to (\mathrm{Del\check{C}}^X)^r$, by the Dowker Nerve Theorem 3.9. The first homotopy equivalence is trivial.

Lemma 5.2. *Let $X \subseteq \mathbb{R}^d$ be a finite set. The inclusion*

$$B_{\mathrm{Alpha}^X}(x, r) \hookrightarrow B_{\check{\mathrm{C}}\mathrm{ech}^X}(x, r)$$

of balls gives us the identity map $(\mathrm{Alpha}^X)^r \to (\check{\mathrm{C}}\mathrm{ech}^X)^r$ *for $r \geq 0$.*

[6] It is technically a simple-homotopy equivalence if we assume X is in general position and we use the gradient collapse argument.

Proof. We get injectivity by hypothesis. For surjectivity, first note that

$$B_{\mathrm{Alpha}^X}(x,r) = B_{\mathrm{Del}^X}(x,r) \cap B_{\check{\mathrm{C}}\mathrm{ech}^X}(x,r).$$

Given $y \in B_{\check{\mathrm{C}}\mathrm{ech}^X}(x,r)$, i.e., $\check{\mathrm{C}}\mathrm{ech}^X(y,x) < r$, there exists some x' in X such that $y \in \mathrm{Vor}(X,x')$. By the definitions of a Voronoi cell and the Čech dissimilarity, we get

$$\begin{aligned}\check{\mathrm{C}}\mathrm{ech}^X(y,x') &= d(y,x')\\ &\le d(y,x)\\ &= \check{\mathrm{C}}\mathrm{ech}^X(y,x) < r.\end{aligned}$$

This means y is also in $B_{\mathrm{Del}^X}(x',r) \cap B_{\check{\mathrm{C}}\mathrm{ech}^X}(x',r)$, as desired. $\square$

Hence, the identity $(\mathrm{Alpha}^X)^r \to (\check{\mathrm{C}}\mathrm{ech}^X)^r$ in Lemma 5.2 gives the first homotopy equivalence

$$\underbrace{|N\mathrm{Alpha}_r^X|}_{|\mathrm{Alpha}_r(X)|} \simeq \underbrace{|N\check{\mathrm{C}}\mathrm{ech}_r^X|}_{|\check{\mathrm{C}}\mathrm{ech}_r(X)|}$$

in the Equivalent Euclidean Data Complexes Theorem 5.1. So we are left with showing that there is a homotopy equivalence $|N\mathrm{Del}\check{\mathrm{C}}_r^X| \simeq |N\check{\mathrm{C}}\mathrm{ech}_r^X|$. We will define a deformation retract using the straight-line homotopy.

Lemma 5.3. *The line segment between any point* $(p,q) \in (\mathrm{Del}\check{\mathrm{C}}^X)^r$ *and* (q,q) *is contained in* $(\mathrm{Del}\check{\mathrm{C}}^X)^r$.

Proof. Having fixed X, the superscript on dissimilarities will be omitted in order to simplify notation. Let $\gamma\colon [0,1] \to \mathbb{R}^d$ be the function $\gamma(s) = (1-s)q + sp$, giving the straight line between p and q in $\mathbb{R}^d$. Then all we need to show is that the point $(\gamma(s),q) \in \mathbb{R}^{2d}$ is in $\mathrm{Del}\check{\mathrm{C}}^r$ for all $s \in [0,1]$. Note that $B_{\check{\mathrm{C}}\mathrm{ech}}(x,r) = B_d(x,r)$ for all $r > 0$.

If $(p,q) \in \mathrm{Del}\check{\mathrm{C}}^r$, then we know there has to be some $x \in X$ such that $p \in B_{\mathrm{Del}}(x,r)$ and $q \in B_d(x,r)$, by definition of $\mathrm{Del}\check{\mathrm{C}}$. The former statement is equivalent to saying $d(p,\mathrm{Vor}(X,x)) < h^-(r)$ since $p \in B_{\mathrm{Del}}(x,r)$ and $\mathrm{Del}_x(p) = h(d(p,\mathrm{Vor}(X,x)))$ is true, if and only if, $h(d(p,\mathrm{Vor}(X,x))) < r$. But by applying h^- and noting that $d(p,\mathrm{Vor}(X,x)) < \infty$, we get that $d(p,\mathrm{Vor}(X,x)) < h^{-1}(r) = h^-(r)$. In particular, pick $p' \in \mathrm{Vor}(X,x)$ such that $d(p,p') < h^-(r)$. We now define a line segment between q and p' given by the function $\gamma'\colon [0,1] \to \mathbb{R}^d$ where $\gamma'(s) = (1-s)q + sp'$. Suppose that $(\gamma'(s),q)$ is in $\mathrm{del}\check{\mathrm{C}}^r$ for some $s \in [0,1]$. Then there exists an $x' \in X$ such that $d(x',q) < t$ and $\gamma'(s) \in \mathrm{Vor}(X,x')$. Moreover, $(\gamma(s),q) \in \mathrm{Del}\check{\mathrm{C}}^r$ since $d(\gamma(s),\gamma'(s)) < h^-(t)$ and $d(\gamma'(s),\mathrm{Vor}(X,x')) = 0$.

What we are left with is to show that $(\gamma'(s), q)$ is in $\mathrm{del\check{C}}^r$ for any $s \in [0,1]$. Hence, since the Voronoi cells cover $\mathbb{R}^d$, we can assume $\gamma'(s)$ is in $\mathrm{Vor}(X, y)$ for some $s \in [0, 1)$ and some $y \in X$, and we will just need to prove that $q \in B_d(y, r)$. We may assume $y \neq x$ without loss of generality. We can cover X using the hyperplane H that separates x and y, i.e.,

$$H = \{z \in X \mid d(x, z) = d(y, z)\}.$$

If we let

$$H_+ = \{z \in X \mid d(x, z) \geq d(y, z)\} \quad \text{and} \quad H_- = \{z \in X \mid d(x, z) \leq d(y, z)\},$$

we have $H_+ \cup H_- = X$ and $H_+ \cap H_- = H$. We see that $\gamma'(s)$ must be in H_+ since it is in $\mathrm{Vor}(X, y)$. Similarly, p' must be in H_- as it is in $\mathrm{Vor}(X, x)$. Since the line segment γ' between q and p' is either contained in H, or traverses H at most once, we must have $q \in H_+$. Recalling our assumption that $q \in B_d(x, r)$, we have $d(y, q) \leq d(x, q) < r$ by definition of H_+, which means $q \in B_d(y, r)$, as desired. $\square$

Corollary 5.4. *The map $(\check{\mathrm{C}}\mathrm{ech}^X)^r \to (\mathrm{Del\check{C}}^X)^r$ given by $p \mapsto (p, p)$, is a deformation retract.*

Since a deformation retract is a particular instance of a homotopy equivalence, Corollary 5.4 and the Dowker Nerve Theorem 3.9 gives the desired homotopy equivalence

$$\underbrace{|N\check{\mathrm{C}}\mathrm{ech}_r^X|}_{|\check{\mathrm{C}}\mathrm{ech}_r(X)|} \simeq \underbrace{|N\mathrm{Del\check{C}}_r^X|}_{|\mathrm{Del\check{C}}_r(X)|}.$$

This finishes the geometric proof of the Equivalent Euclidean Data Theorem 5.1.

5.2 Gradient Collapse Argument

We now turn to Bauer and Edelsbrunner's [3] slightly stronger version of the Equivalent Euclidean Data Complexes Theorem 5.1, using generalized discrete Morse theory. We do, however, need more work to get there than we did with the more direct proof above. Thus, we start with some lemmas before eventually getting to the simplicial collapse of the three Euclidean data complexes, i.e., we show that they are simple-homotopy equivalent if X is in general position. In particular, we will demonstrate two collapses of Euclidean data complexes:

$$\check{\mathrm{C}}\mathrm{ech}_r(X) \searrow \mathrm{Del\check{C}}_r(X) \searrow \mathrm{Alpha}_r(X).$$

As we saw in Section 2, all three of these are instances of the selective Delaunay-Alpha complex. We will, hence, show the two collapses as particular instances of a collapse of selective Delaunay-Alpha complexes by choosing appropriate subsets $E, F \subseteq X$.

5.2.1 Preliminary Lemmas

We will first need to cover some necessary lemmas for us to prove the two gradient collapses above. Given a simplex $\sigma \in K$ and a vertex x belonging to a simplicial complex K, we follow the notation in [3] by writing $\sigma - x$ instead of $\sigma - \{x\}$ and $\sigma + x$ instead of $\sigma \cup \{x\}$. We have decided to only provide references for most of the lemmas, as their statements are more relevant to us than the technical proofs themselves. First, we want different ways of constructing discrete gradients for a simplicial complex.

Lemma 5.5 (Vertex Gradient Lemma)**.** *Let K be a simplicial complex, V a discrete vector field on K and x a vertex of K. If every pair in V is of the form $\{\sigma - x, \sigma + x\}$ for some simplex $\sigma \in K$, then V is a discrete gradient on K.*

Proof. See Lemma 5.1 in [3]. □

By choosing some vertex x in our simplicial complex, the Vertex Gradient Lemma 5.5 gives us a useful class of discrete gradients. Clearly, σ is either equal to $\sigma - x$ or $\sigma + x$, which means we get a discrete gradient V where every non-critical simplex σ is either in a pair $\{\sigma, \sigma + x\}$ or $\{\sigma - x, \sigma\}$ in V. We can also construct a discrete gradient for a simplicial complex by composing the discrete gradient of a subcomplex with one defined on its complement.

Lemma 5.6 (Gradient Composition Lemma)**.** *Let L be a subcomplex of the simplicial complex K with discrete gradients V_L and V_K, respectively. If every pair in V_K is disjoint from L, then the pairs in $V_K \cup V_L$ define a discrete gradient on K.*

Proof. See Lemma 4.3 in [22]. □

One can, for instance, take the union $K \cup L$ of two disjoint simplicial complexes K and L, with discrete gradients V_K and V_L, respectively. Then V_K is a discrete gradient for $K \cup L$ and is disjoint from L. Hence, the Gradient Composition Lemma 5.6 tells us $V_K \cup V_L$ is a discrete gradient on $K \cup L$. In effect, the lemma reduces the number of critical simplices. The composed gradient is more descriptive of the simplicial complex we are working with by potentially collapsing the original one to an even smaller subcomplex than before the composition, if we use the Gradient Collapse Theorem 4.12.

We can also describe the generalized discrete gradient when taking the sum of two generalized discrete Morse functions. It is a refinement because the sum of two functions on two different simplicial complexes has to be restricted to their intersection.

Lemma 5.7 (Sum Refinement Lemma). *Let K and L be simplicial complexes with generalized discrete Morse functions $f\colon K \to \mathbb{R}$ and $g\colon L \to \mathbb{R}$. Moreover, let their generalized discrete gradients be W_f and W_g. Then the sum of functions $f+g\colon K \cap L \to \mathbb{R}$ is a generalized discrete Morse function with generalized discrete gradient*

$$W_{f+g} = \{I \cap J \mid I \in W_f, J \in W_g \text{ and } I \cap J \neq \emptyset\}.$$

Next, we consider when $S(\sigma, E)$ remains unchanged while removing a point from σ or E.

Lemma 5.8 (Same Sphere Lemma). *Given a finite set $X \subseteq \mathbb{R}^d$, let E be a subset of X with σ a simplex in $\mathrm{Del}(X, E)$, and let $S = S(\sigma, E)$. Then*

(1) $S(\sigma - x, E) = S(\sigma + x, E)$ if and only if $x \in \mathrm{Incl}\, S - \mathrm{Front}\, S$, and

(2) $S(\sigma, E - y) = S(\sigma, E + y)$ if and only if $y \in \mathrm{Excl}\, S - \mathrm{Back}\, S$.

Proof. For (1), recall that S must be equal to $S(\sigma - x, E)$ or $S(\sigma + x, E)$. Moreover, we get that $\sigma \subseteq \mathrm{Incl}\, S$ since S includes σ, while the Combinatorial KKT Theorem 4.18 gives $\mathrm{Front}\, S \subseteq \sigma$. This implies that x being in $\mathrm{Incl}\, S - \mathrm{Front}\, S$ is equivalent to $\mathrm{Front}\, S \subseteq \sigma - x$ and $\sigma + x \subseteq \mathrm{Incl}\, S$, which again is equivalent to $\sigma - x$ and $\sigma + x$ belonging to $[\mathrm{Front}\, S, \mathrm{Incl}\, S]$ since $\sigma - x$ is contained in $\sigma + x$. This gives the desired result. The proof of (2) is similar. □

Using the Same Sphere Lemma 5.8, we get the most crucial lemmas for obtaining the desired gradient collapses. The next two will be more specific about which point can be added or removed from a simplex without changing the smallest circumscribed sphere of said simplex, given different criteria.

Lemma 5.9 (First Simplex Pairing Lemma). *Given a finite set $X \subseteq \mathbb{R}^d$, let $E \subseteq F$ be subsets of X with σ a simplex in $\mathrm{Del}(X, F)$ such that $S(\sigma, E) \neq S(\sigma, F)$. Then there exists a point $x \in F - E$ such that*

(1) $S(\sigma - x, E) = S(\sigma + x, E)$ and

(2) $S(\sigma - x, F) = S(\sigma + x, F)$.

Proof. See Lemma 5.5 in [3]. □

Define A to be equal to $\mathrm{Del}(X, E) - \mathrm{Del}(X, F)$.

Lemma 5.10 (Second Simplex Pairing Lemma). *Given a finite set $X \subseteq \mathbb{R}^d$, let $E \subseteq F$ be subsets of X, and let σ be in A. Then there exists a point $x \in F - E$ such that*

(1) $S(\sigma - x, E) = S(\sigma + x, E)$ and

(2) neither $\sigma - x$, nor $\sigma + x$, are in $\mathrm{Del}(X, F)$.

Proof. See Lemma 5.7 in [3]. $\square$

The final lemma we will need tells us something about the ability to make a consistent choice of x in the Second Simplex Pairing Lemma 5.10. We can, namely, change $\sigma - x$ or $\sigma + x$ for σ without it changing the result.

Lemma 5.11 (Second Consistent Pairing Lemma)**.** *Given a finite set $X \subseteq \mathbb{R}^d$, let $E \subsetneq F$ be subsets of X. Then there is a map*

$$\psi \colon A \to F - E$$

such that, for a given $\sigma \in A$, the point $x = \psi(\sigma)$ satisfies the properties of Lemma 5.10 and is equal to both $\psi(\sigma - x)$ and $\psi(\sigma + x)$.

Proof. See Lemma 5.8 in [3]. $\square$

5.2.2 Collapsing Theorem

Having all the lemmas from the previous section, we are now ready to prove the stronger version of the Equivalent Euclidean Data Complexes Theorem 5.1. It is stronger because it gives us a simple-homotopy, but we do need to assume X is in general position. Note that this is Theorem 5.9 in [3] but in terms of the new selective Delaunay-Alpha complex. Hence, our new complex is simple-homotopy equivalent to the selective Alpha complex. As was mentioned in Section 4, we know two simplicial complexes are simple-homotopy equivalent if there exists a simplicial collapse from one to the other.

Theorem 5.12 (Selective Alpha Collapsing Theorem)**.** *Let $X \subseteq \mathbb{R}^d$ be a finite set in general position with subsets $E \subseteq F$. Then*

$$\mathrm{Alpha}_r(X, E) \searrow \mathrm{DelA}_r(X, F, E) \searrow \mathrm{Alpha}_r(X, F)$$

for $r \geq 0$.

Proof. We can get both collapses if we construct appropriate discrete gradients and use the two gradient collapsing theorems. This will be done with the help of the preliminary lemmas from the previous subsection.

We begin with the last collapse. First, note that

$$\mathrm{DelA}_r(X, F, E) = \mathrm{Del}(X, F) \cap \mathrm{Alpha}_r(X, E).$$

Also recall that the radius functions s_E and s_F are generalized discrete Morse functions by Theorem 4.19. Let W_E and W_F be their respective generalized discrete gradients. Then $s_E + s_F\colon \mathrm{Del}(X,F) \to \mathbb{R}$ is also a generalized discrete Morse function with generalized discrete gradient

$$W_{E+F} = \{I \cap J \mid I \in W_E, J \in W_F \text{ and } I \cap J \neq \emptyset\},$$

by the Sum Refinement Lemma 5.7, and since

$$\mathrm{Del}(X,E) \cap \mathrm{Del}(X,F) = \mathrm{Del}(X,F)$$

whenever $E \subseteq F$, by Lemma 2.15. A simplex σ that is in the intersection of $\mathrm{Alpha}_r(X,E)$ and $\mathrm{Del}(X,F)$, but is not a simplex in $\mathrm{Alpha}_r(X,F)$, gives a sphere $S(\sigma,E)$ with radius at most r, while $S(\sigma,F)$ must have a radius larger than r. A subset U of the intervals in W_{E+F} gives a partition of all such simplices. Moreover, we know that $S(\sigma,E) \neq S(\sigma,F)$ since their radii are different, which implies that U does not contain any singular intervals, by the First Simplex Lemma 5.9. Then the Generalized Gradient Collapsing Theorem 4.16 gives us the desired collapse

$$\mathrm{DelA}_r(X,F,E) \searrow \mathrm{Alpha}_r(X,F).$$

We get the first collapse as a particular case of $\mathrm{Del}(X,E) \searrow \mathrm{Del}(X,F)$ since

$$\mathrm{Alpha}_r(X,E) \subseteq \mathrm{Del}(X,E) \quad \text{and} \quad \mathrm{DelA}_r(X,F,E) \subseteq \mathrm{Del}(X,F)$$

whenever $E \subseteq F$ and $r \geq 0$. Hence, we will partition $\mathrm{Del}(X,E) - \mathrm{Del}(X,F)$ into pairs using the Second Simplex Pairing Lemma 5.10. As in the proof of the Second Consistent Pairing Lemma 5.11, we fix a total ordering $x_1, x_2, \ldots, x_n$ on the points in X and recursively define a filtration of $\mathrm{Del}(X,E)$. Let

$$K_0 = \mathrm{Del}(X,E)$$

and

$$K_i = K_{i-1} - \{\sigma \subseteq X \mid \psi(\sigma) = x_i\}$$

for $i = 1, 2, \ldots, n$. The resulting filtration is thus

$$\mathrm{Del}(X,F) = K_n \subseteq K_{n-1} \subseteq \cdots \subseteq K_0 = \mathrm{Del}(X,E).$$

Similarly, we define the sets

$$P_i = \{\{\sigma - x_i, \sigma + x_i\} \mid \sigma \subseteq X \text{ and } \psi(\sigma) = x_i\}$$

containing pairs for $i = 1, 2, \ldots, n$. Each P_i gives rise to a discrete gradient V_i that partitions

$$K_{i-1} - K_i = \{\sigma \subseteq X \mid \psi(\sigma) = x_i\}.$$

By construction, V_i consists only of pairs, which implies $K_{i-1} \searrow K_i$, by the Gradient Collapsing Theorem 4.12. Then the union $\bigcup_{i=1}^n P_i$ gives a partition of $\mathrm{Del}(X, E) - \mathrm{Del}(X, F)$ into pairs, and we may apply the Gradient Composition Lemma 5.6 inductively for our V_i and K_i. By again applying the Gradient Collapsing Theorem 4.12, this gives us a discrete gradient on $\mathrm{Del}(X, E)$ that defines a collapse $\mathrm{Del}(X, E) \searrow \mathrm{Del}(X, F)$. $\square$

Corollary 5.13. *Let $X \subseteq \mathbb{R}^d$ be a finite set in general position with subsets $E \subseteq F$. Then*

$$\check{\mathrm{C}}\mathrm{ech}_r(X) \searrow \mathrm{Del}\check{\mathrm{C}}_r(X) \searrow \mathrm{Alpha}_r(X)$$

for $r \geq 0$.

Proof. If we let $E = \emptyset$ and $F = X$, then the Selective Alpha Collapsing Theorem 5.12 gives the desired collapse since $\mathrm{Alpha}_r(X, \emptyset)$ and $\mathrm{Alpha}_r(X, X)$ are equal to $\check{\mathrm{C}}\mathrm{ech}_r(X)$ and $\mathrm{Alpha}_r(X)$, respectively. Moreover, we know

$$\mathrm{DelA}_r(X, X, \emptyset) = \mathrm{Del}\check{\mathrm{C}}_r(X),$$

by Table 1. $\square$

This concludes the second proof of the Equivalent Euclidean Data Complexes Theorem 5.1. Since the selective Delaunay-Alpha dissimilarity generalizes all the Euclidean data complexes in this thesis, we can rewrite the Selective Alpha Collapsing Theorem 5.12 as

$$\mathrm{DelA}_r(X, E, E) \searrow \mathrm{DelA}_r(X, F, E) \searrow \mathrm{DelA}_r(X, F, F),$$

while Corollary 5.13 becomes

$$\mathrm{DelA}_r(X, \emptyset, \emptyset) \searrow \mathrm{DelA}_r(X, F, E) \searrow \mathrm{DelA}_r(X, X, X).$$

We end this section by proving a new corollary. It tells us that, more than generalizing many Euclidean data complexes, the selective Delaunay-Alpha complex is simple-homotopy equivalent to them as well.

Corollary 5.14. *Let $X \subseteq \mathbb{R}^d$ be a finite set in general position with subsets $E \subseteq F$. Then*

$$\check{\mathrm{C}}\mathrm{ech}_r(X) \simeq \mathrm{DelA}_r(X, F, E) \simeq \mathrm{Alpha}_r(X)$$

are two simple-homotopy equivalences for $r \geq 0$.

Proof. If we let $E = \emptyset$ in the Selective Alpha Collapsing Theorem 5.12, then we get the simplicial collapse

$$\check{\mathrm{C}}\mathrm{ech}_r(X) \searrow \mathrm{Alpha}_r(X, F).$$

Hence, the selective Alpha and Čech complexes are simple-homotopy equivalent. However, the same theorem tells us that the selective Alpha complex is simple-homotopy equivalent to the selective Delaunay-Alpha complex, by the theorem's second collapse. Transitivity gives us that the Čech and selective Delaunay-Alpha complexes are simple-homotopy equivalent. We get the second simple-homotopy equivalence if we let $F = X$ and use a similar argument. $\square$

6 Conclusion

In this thesis, we have considered a range of Euclidean data complexes, with the Čech and Alpha complex being the most important. We have seen that the former is quite descriptive in that it is homotopy equivalent to the union of balls centered at each data point. However, the latter is preferable due to computational efficiency and it having low-dimensional simplices that locally resemble the underlying manifold better. An important goal was to show that we can work with the Alpha complex without sacrificing much of the topological content in the Čech complex. In the process, we proved that the Delaunay-Čech complex is equivalent to them both using the geometric and gradient collapse argument.

We also introduced a new Euclidean data complex that we called the selective Delaunay-Alpha complex. We saw how it generalizes all the other complexes, and that it fits into a simplicial collapse in the Selective Alpha Collapsing Theorem 5.12. Most importantly, we showed that it is simple-homotopy equivalent to the Alpha and Čech complex, and that, as a result, it is also equivalent to all the other Euclidean data complexes in Section 2. Hence, we can freely choose which Euclidean data complex to work with, knowing it will be the same simplicial complex up to homotopy equivalence. In particular, filtrations of the Čech, Alpha, Delaunay-Čech and selective Delaunay-Alpha complex have isomorphic persistent homology [3, p. 18].

Before discussing potential further research, we will present a failed attempt at proving Corollary 5.14 using elementary geometry, similar to what was done in Lemma 5.3. It is not in general true that the line segment between two points (q, q) and (p, q) in $(\mathrm{DelA}^X_{E,F})^r$ is contained in $(\mathrm{DelA}^X_{E,F})^r$. Suppose, on the contrary, that such a line segment does exist. We will suppress sub- and superscripts on all dissimilarities from now on. If $(p, q) \in \mathrm{DelA}^r$, then we know there has to be some $x \in X$ such that $p \in B_{\mathrm{Del}}(x, r)$ and $q \in B_{\mathrm{Alpha}}(x, r)$, by definition of DelA. Recall that $p \in B_{\mathrm{Del}^X_E}(x, r)$ is equivalent to $p \in \mathrm{Vor}(E, x)$, and that in the proof of Lemma 5.3 the separating plane H was useful because we knew the Voronoi cell of x gives all the points lying closer to x than any other point $y \in X$. This is no longer true when $E \neq X$.

An attempt at remedying this problem could have been to find an appropriate $e \in E$ such that x and p are both in $\mathrm{Vor}(E, e)$. Choose an $e \in E$ such that $d(e, x) \leq d(e', x)$ for all $e' \in E$. Now it is true that x is in $\mathrm{Vor}(E, e)$, but not necessarily true that p is. See Figure 19 for a counter-example. Thus, this seems to imply that the choice of $E \subseteq X$ matters for the geometry of the selective Delaunay-Alpha complex. For future research, one might attempt a similar proof while being more careful about how the choice of $E, F \subseteq X$ impact the geometry of $\mathrm{DelA}_r(X, E, F)$. Perhaps a better approach in obtaining a geometric proof

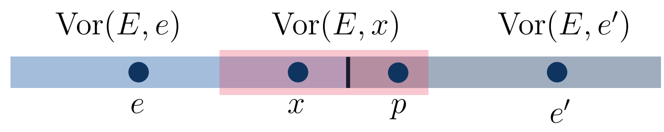

Figure 19: When $x \notin E$ it is possible for $p \notin \mathrm{Vor}(E, e)$ even if $p \in \mathrm{Vor}(E, x)$.

would be to do as in Corollary 5.14, i.e., we could try to show that the selective Delaunay-Alpha complex is homotopy equivalent to the selective Alpha complex first. Does assuming X is in general position also play a role? Can we obtain a homotopy equivalence without this assumption?

However, having Corollary 5.14 is more than sufficient to make the selective Delaunay-Alpha complex interesting. We think our new complex might be useful due to the increased number of parameters giving more control to create a specific type of complex. The selective Delaunay-Alpha complex can, namely, produce Euclidean data complexes that are not equal to any of the others, even if it is simple-homotopy equivalent to all of them. This could possibly lead to the use of machine learning methods to optimize these parameters and find the complex that would best suit one's needs, e.g., using regularization to find the optimal dimension for our simplices. In other words, there are many avenues – in both mathematics and computer science – to explore related to the selective Delaunay-Alpha complex.

References

[1] Paul Bendich Abraham D. Smith and John Harer. "Persistent obstruction theory for a model category of measures with applications to data merging". In: *Transactions of the American Mathematical Society Series B* 3.1-2 (2021), pp. 1–38. DOI: https://doi.org/10.1090/btran/56.

[2] Franz Aurenhammer. "Voronoi diagrams – a survey of a fundamental geometric data structure". In: *ACM Compututing Surveys* 23.3 (1991), pp. 345–405. DOI: https://doi.org/10.1145/116873.116880.

[3] Ulrich Bauer and Herbert Edelsbrunner. "The Morse theory of Čech and Delaunay complexes". In: *Transactions of the American Mathematical Society* 369.5 (2017), pp. 3741–3762. DOI: https://doi.org/10.1090/tran/6991.

[4] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer-Verlag New York, 2006.

[5] Nello Blaser and Morten Brun. "Relative persistent homology". In: *36th International Symposium on Computational Geometry (SoCG 2020)*. Vol. 164. 2020, 18:1–18:10. DOI: https://doi.org/10.4230/LIPIcs.SoCG.2020.18.

[6] Morten Brun and Nello Blaser. "Sparse Dowker nerves". In: *Journal of Applied and Computational Topology* 3.1-2 (2019), pp. 1–28. DOI: https://doi.org/10.1007/s41468-019-00028-9.

[7] Gunnar Carlsson. "Topology and data". In: *Bulletin of the American Mathematical Society* (2009), pp. 255–308. DOI: https://doi.org/10.1090/S0273-0979-09-01249-X.

[8] Gunnar Carlsson et al. "On the local behavior of spaces of natural images". In: *New Results and New Trends in Computer Science* 76 (2008). Ed. by Hermann Maurer, pp. 1–12. DOI: https://doi.org/10.1007/s11263-007-0056-x.

[9] Joseph Minhow Chan, Gunnar Carlsson, and Raul Rabadan. "Topology of viral evolution". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.46 (2013), pp. 18566–18571. DOI: https://dx.doi.org/10.1073%2Fpnas.1313480110.

[10] Manoj K. Chari. "On discrete Morse functions and combinatorial decompositions". In: *Discrete Mathematics* 217.1 (2000), pp. 101–113. DOI: https://doi.org/10.1016/S0012-365X(99)00258-7.

[11] Marshall M. Cohen. *A Course in Simple-Homotopy Theory*. Vol. 10 of Graduate Texts in Mathematics. Springer-Verlag New York, 1973. DOI: https://doi.org/10.1007/978-1-4684-9372-6.

[12] Albrecht Dold. *Lectures on Algebraic Topology*. Classics in Mathematics. Springer-Verlag Berlin Heidelberg, 1995.

[13] Clifford H. Dowker. "Homology groups of relations". In: *Annals of mathematics* (1952), pp. 84–95. DOI: https://doi.org/10.2307/1969768.

[14] Ludovic Duponchel. "When remote sensing meets topological data analysis". In: *Journal of Spectral Imaging* 7.1 (2018), a1. DOI: https://doi.org/10.1255/jsi.2018.a1.

[15] Herbert Edelsbrunner and John Harer. *Computational Topology: An Introduction*. American Mathematical Society, 2010.

[16] Paul Embrechts and Marius Hofert. *A note on generalized inverses*. Feb. 2014. URL: https://people.math.ethz.ch/~embrecht/ftp/generalized_inverse.pdf (visited on 10/11/2020).

[17] Kaspar Fischer, Bernd Gärtner, and Martin Kutz. "Fast smallest-enclosing-ball computation in high dimensions". In: *European Symposium on Algorithms*. Springer. 2003, pp. 630–641. DOI: https://doi.org/10.1007/978-3-540-39658-1_57.

[18] Robin Forman. "A user's guide to discrete Morse theory". In: *S´eminaire Lotharingien de Combinatoire* 48 (Dec. 2001).

[19] Robin Forman. "Morse theory for cell complexes". In: *Advances in Mathematics* 134.1 (1998), pp. 90–145. DOI: https://doi.org/10.1006/aima.1997.1650.

[20] Ragnar Freij. "Equivariant discrete Morse theory". In: *Discrete Mathematics* 309.12 (2009), pp. 3821–3829. DOI: https://doi.org/10.1016/j.disc.2008.10.029.

[21] Robert Ghrist. *Elementary Applied Topology*. 10th ed. Createspace, 2014.

[22] Jakob Jonsson. *Simplicial Complexes of Graphs*. Vol. 28. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2008. DOI: https://doi.org/10.1007/978-3-540-75859-4.

[23] Christopher Olah. *Neural networks, manifolds, and topology*. 2014. URL: https://colah.github.io/posts/2014-03-NN-Manifolds-Topology/ (visited on 05/28/2021).

[24] Walter Rudin. *Principles of Mathematical Analysis*. Vol. 3. McGraw-Hill New York, 1964.

[25] Emo Welzl. "Smallest enclosing disks (balls and ellipsoids)". In: *New Results and New Trends in Computer Science* (1991). Ed. by Hermann Maurer, pp. 359–370. DOI: https://doi.org/10.1007/BFb0038202.