Mesh Generation and Geometric Persistent Homology

Donald R. Sheehy

CMU-CS-11-121

July 2011

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Gary L. Miller, Chair
Guy Blelloch
Daniel D.K. Sleator
Afra J. Zomorodian, Dartmouth College

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

This research was sponsored by the National Science Foundation under grant numbers CCF-0635257 and CCF-1065106. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of any sponsoring institution, the U.S. government or any other entity.

Keywords: Computational Geometry, Mesh Generation, Persistent Homology

Abstract

Mesh generation is a tool for discretizing functions by discretizing space. Traditionally, meshes are used in scientific computing for finite element analysis. Algorithmic ideas from mesh generation can also be applied to data analysis.

Data sets often have an intrinsic geometric and topological structure. The goal of many problems in geometric inference is to expose this intrinsic structure. One important structure of a point cloud is its **geometric persistent homology**, a multi-scale description of the topological features of the data with respect to distances in the ambient space.

In this thesis, I bring tools from mesh generation to bear on geometric persistent homology by using a mesh to approximate distance functions induced by a point cloud. Meshes provide an efficient way to compute geometric persistent homology. I present the first time-optimal algorithm for computing quality meshes in any dimension. Then, I show how these meshes can be used to provide a substantial speedup over existing methods for computing the full geometric persistence information for range of distance functions.

Acknowledgements

There are many people without whom this thesis would not be possible. Most directly, I would like to thank my coauthors on the various papers comprising this thesis: Benoît Hudson, Steve Oudot, Todd Phillips, and Gary Miller. I am especially thankful for the work that Benoît did in implementing a 4-dimensional mesh generator in order to test our mesh-based persistence method on a real example. Steve was tremendously helpful in keeping me honest when that ever-present specter of topological intuition threatened to lead me into nonrigorous proofs. He also helped me secure a postdoc position that allowed me to write this thesis without the added stress of an academic job search. I owe a special debt of gratitude to Gary and Todd as my closest collaborators throughout my graduate studies. As anyone in research knows, posing good problems is as important as solving them. I often found my work directed towards proving their insightful conjectures.

As my PhD advisor, Gary Miller has been a great help in my development into research maturity. His emphasis on clear intuition greatly improved this thesis. The other members of my thesis committee, Guy Blelloch, Danny Sleator, and Afra Zomorodian were also very helpful in the preparation of the final draft. I would like to call special attention to Afra, whose careful reading and copious notes were invaluable during the editing process.

I am also very grateful for the early direction and guidance I received in my undergraduate days. Bernard Chazelle first inspired in me a great love for computational geometry. Satyan Devadoss showed me a model of research and life that I still aspire to. Moreover, he recognized my interests and first handed me a copy of Edelsbunner's *Geometry and Topology for Mesh Generation*, perhaps the single most formative book in my academic development.

Completing a dissertation is not easy (at least it wasn't for me). I don't think it would be possible without good company outside of research activities. For that, I am grateful to the Lunch Bunch (my twice weekly basketball game) and the Movie Night crowd (my weekly dinner and movie group).

I would also like to thank my family for their support: my parents Don and Kathy, and my sisters Colleen and Cristin (my biggest fan).

And finally I thank my wife, Julia, for keeping me fed, clothed, and sane.

Contents

Chapter 1

Introduction

1.1 Mesh Generation

The goal of mesh generation is to decompose a geometric domain into simple elements. For example, the square on the left in Figure 1.1 is decomposed into triangles. A discrete description of a space allows for discrete approximations to functions on that space. Thus, mesh generation is ubiquitous in physical simulation where it allows for the numerical solution to partial differential equations.

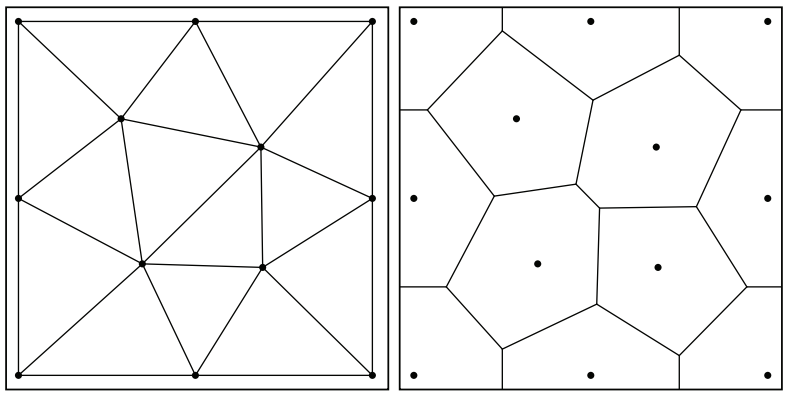

Figure 1.1: **Left:** The Delaunay triangulation of a set of points. The triangulation decomposes the convex closure of the points. **Right:** The Voronoi diagram of the same set of points. The diagram decomposes all of $\mathbb{R}^2$.

The decompositions in this thesis all come from the Delaunay triangulation or its dual, the Voronoi diagram. Roughly speaking, the Delaunay triangulation of a set of points $P \subset \mathbb{R}^2$ has a triangle for every 3 points whose circumscribing circle contains no other points of P. The Voronoi diagram is dual to the Delaunay triangulation. It decomposes the plane into polygons called Voronoi cells, one for each $q \in P$, such that the Voronoi cell of q is the set of all points whose nearest neighbor in P is q. Figure 1.1 shows a simple example of a Delaunay triangulation and a Voronoi diagram on the same set of points. There is a natural way to define Delaunay triangulations and Voronoi diagrams in higher dimensions (see Chapter 2). One advantage of using the Delaunay triangulation

for mesh generation is that it gives a unique decomposition of space defined only by the set of vertices. Thus, we will speak primarily about point sets and the decomposition will be secondary, used mainly for definitions and proofs.

Let P be a set of n points in $\mathbb{R}^d$. A mesh generation algorithm will output a superset M of P such that the Voronoi cells have aspect ratio at least some constant. That is, the cells (modulo some on the boundary) should all be sufficiently "round". The set M along with the Delaunay triangulation of M is called the **mesh**. The extra points added are called **Steiner points**. The **size** of the mesh is its number of points, $|M|$. This is to be distinguished from the **complexity** of Vor_M, which is the number of cells in the Voronoi diagram of M.

We will focus on three main goals of mesh generation:

1. The mesh should **conform** to a set of input points. That is, every input point should appear as a vertex in the output.
2. The mesh should satisfy some **quality** condition. That is, every cell in the output should be geometrically "nice", where several different metrics of quality are used in practice.
3. The mesh should have **optimal size**. Up to constant factors, no other conforming, quality mesh can have fewer vertices.

For many meshing applications, the conforming condition is extended to also include higher order features such as edges, faces, or even piecewise-smooth complexes. In this thesis, we limit the input to points, which is most relevant when using meshes to do data analysis; the input points are the data.

There are several metrics used to measure the quality of a mesh. For example, in 2D triangulation, the measure of the smallest angle is commonly used. The idea behind these quality metrics is to give a purely geometric condition that can give guarantees about how well a mesh will work for a given application. For example, the quality a mesh influences the quality of a solution in finite element analysis [SF73]. Much work has been done on mesh quality measures (see [She02b, BA76, MTTW95, MTTW99] for discussion on mesh quality conditions relevant to Delaunay meshing). For triangulations and other simplicial meshes, the ratio of the circumradius to the shortest edge may be used in place of the angle condition (see Figure 1.2). A similar notion of quality may be defined on the Voronoi diagram. It is called the aspect ratio of the Voronoi cells and is illustrated in Figure 1.3 (see page 9 for a formal definition).

These quality conditions may leave some simplices called slivers which are known to be problematic in physical simulation. However, there are many theoretical and practical approaches to dealing with slivers in meshes [CDE$^+$00, ELM$^+$00, Li00, LT01, EG02, Li03, Lab06]. Despite this wealth of research directed at removing slivers, it remains a major research problem in mesh generation.

Figure 1.4 shows an example of a set of points, its Voronoi diagram, and the Voronoi diagram after Steiner points were added. Note that the Voronoi cells of the input points are long and skinny, whereas the output Voronoi cells are round.

The third goal of size optimality requires lower bounds. To prove these lower bounds, we use the **Ruppert local feature size** $\mathbf{f}_P : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$, defined as the distance to the *second* nearest

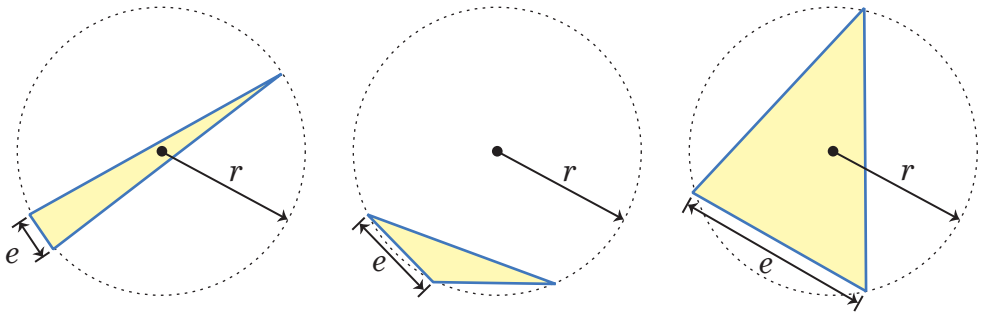

Figure 1.2: The ratio of the radius of a triangle to the length of its shortest edge is a commonly used measure of mesh quality.

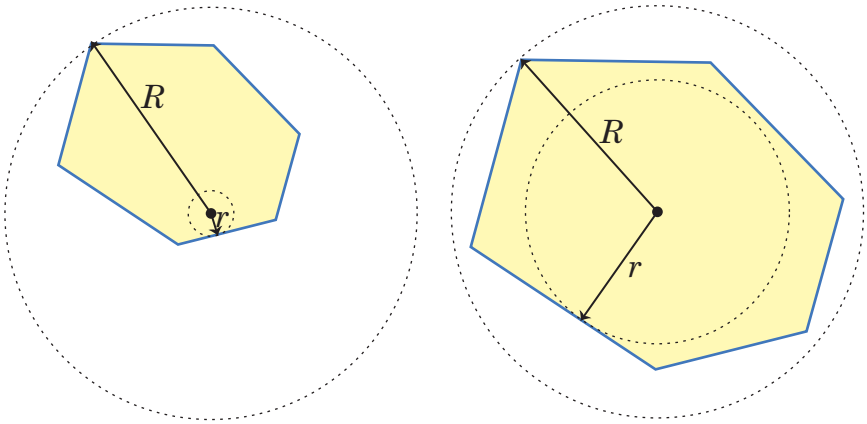

Figure 1.3: The aspect ratio of a Voronoi cell is illustrated with two examples. Note that although the Voronoi cells are geometrically similar, they have different aspect ratios because the vertices are not in the same relative location within the cells.

point of P. Ruppert's seminal result on the analysis of 2-dimensional Delaunay refinement reduces the problem of bounding the mesh size to a geometric problem of bounding the integral of a certain measure over the input domain Ω [Rup95]. His method generalizes naturally to $\mathbb{R}^d$ to give the following.

$$|M| = \Theta\left(\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d}\right).$$

Note that the statement here contains a big-Θ and not a big-O. Consequently, the analysis doesn't require computing this integral, we only need the upper bound because the matching lower bound guarantees the result will be "optimal". This may seem a bit strange with respect to traditional analysis of algorithms. While it does yield a guarantee of "optimality", it does not yield a simple description of the asymptotic complexity as a function of n. That is, we don't get a clear explanation of what optimal *means*. In fact, the bound cannot be stated directly as a function of n because the integral depends on geometric properties of the input P and the input domain Ω.

We could try to add another parameter to capture this geometric structure of the input. For example, letting Δ denote the **spread** of the input (the ratio of the largest to smallest pairwise

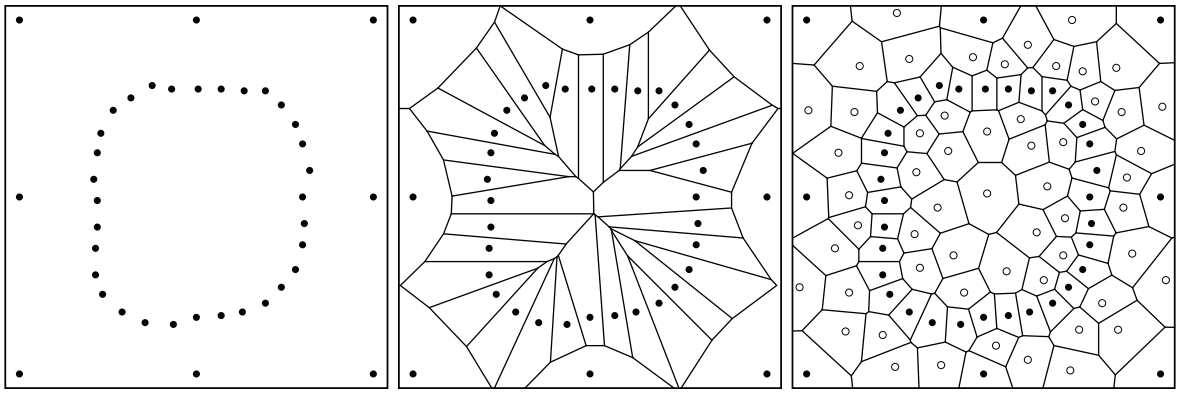

Figure 1.4: An example of point meshing.

distances), it is not difficult to prove that

$$|M| = O(n \log \Delta).$$

This is a conceptual improvement in that it may be easier to think about (and write down) such a bound, but it adds a good deal of slack. There are simple examples such that $|M| = O(n)$, and yet $n \log \Delta = n^2$.

In Chapter 3, we present an approach that gives a tighter bound. Consider any ordering on the points of P such that p_1 and p_2 are the farthest pair and let P_i be the ith **prefix** of this ordering, i.e. $P_i = \{p_1, \ldots, p_i\}$. Let $\theta_i = \frac{\mathbf{f}_{P_i}(p_i)}{\mathbf{f}_{P_{i-1}}(p_i)}$ be the ratio of the distances to the second nearest and nearest neighbor of p_i among the first $i-1$ points. Assuming a reasonably sized bounding domain around P, we will prove that the size of an optimal mesh is

$$|M| = \Theta\left(\sum_{i=3}^{n}(1 + \log \frac{1}{\theta_i})\right).$$

That is, the contribution of each point is the change in the log of its feature size before and after insertion. The fraction θ_i is close to 1 when p_i is roughly equidistant from its nearest and second nearest neighbors among the first $i-1$ points. We say p_i is θ**-medial** if θ_i is larger than some constant θ. When θ is understood, we just say p_i is medial. If there exists an input ordering so that each p_i is medial then the output size will be $O(n)$.

The new bound is asymptotically tight. It also gives an important new insight into the way that we analyze mesh sizing. It reduces the analysis to finding an ordering. For example, to show that an output mesh has linear size, it suffices to find a **well-paced** ordering, i.e. one such that each p_i is medial. Such an ordering is not guaranteed to exist.

Sometimes, it is possible to turn an analytic tool into an algorithm. For example, one could hope to build a meshing algorithm that works by explicitly or implicitly finding an ordering that maximizes θ_i. It turns out that this is exactly what happens in the Sparse Voronoi Refinement

(SVR) algorithm of Hudson et al [HMP06, HMP07]. In SVR, the mesh is constructed incrementally. An input point is inserted only if it is medial. Otherwise, a Steiner point is added. The running time of SVR is $O(n \log \Delta + |M|)$. Thus, we see a return of the dependence on the spread as a result of the point location cost.

We eliminate the dependence on the spread in Chapter 4 where we introduce a method for ordering the input points so the total work of point location is only $O(n \log n)$. We replace the usual quality condition on the aspect ratio of Voronoi cells with a notion of hierarchical quality in which the domain is partitioned into a hierarchical tree of sets so that although individual Voronoi cells may have bad aspect ratio, the union of all the cells in a subtree does have good aspect ratio. This insight strengthens the analogy of Voronoi refinement mesh generation with the closely related class of algorithms that use compressed quadtrees. The result is NetMesh, the first algorithm to achieve optimal $O(n \log n + |M|)$ running time for conforming mesh generation of point sets. Moreover, NetMesh can return a hierarchical quality mesh of size $O(n)$ in $O(n \log n)$ time.

The Ruppert lower bound required us to leave the class of bounded aspect ratio Voronoi diagrams in order to guarantee a linear size output on all inputs. In Chapter 5, we explore how a slight change in the definition of a quality mesh can circumvent these lower bounds. The NetMesh algorithm can be modified to produce Voronoi diagrams in which every cell is **fat**, the ratio of the radii of smallest containing ball and the largest contained ball is bounded by some constant. This is very close to the usual definition of quality for Voronoi diagrams; it only relaxes the requirement that the aspect ratio be measured with respect to balls centered at the point generating the Voronoi cell.

We prove that this slight change in the definition results in some very nice consequences. For example, the Ruppert lower bound can be completely overcome; we prove a linear upper bound on the size of a fat mesh. The full complexity of a fat Voronoi diagram, the sum of the numbers of faces for each dimension, is only linear in the number of vertices. The main focus of Chapter 5 is on bounding the complexity of individual cells of a fat Voronoi diagram. Thus, the average case complexity results known for general fat complexes can sometimes be improved to worst case results when the complex has the special structure of a Voronoi diagram. From an algorithmic perspective, the worst case bound is important because local operations updating a Voronoi diagram will have complexity that depends on the local complexity.

1.2 Geometric Persistent Homology

The second part of this thesis addresses a new application area for Voronoi refinement meshing techniques: topological data analysis. The relatively new field of topological data analysis, attempts to extract topological information from unstructured points, usually assumed to lie in a metric (or geometric) space. It has been applied successfully to many problem domains, including image analysis[CIdSZ08], biology[SMI+08, CBK09], and sensor networks[dSG07b, dSG07a]. See also the survey by Carlsson for background on the topological view of data [Car09].

One of the most widely used techniques of topological data analysis is persistent homology [ELZ02, ZC05]. We address the question of how to approximate the geometric persistent homology of a point set in $\mathbb{R}^d$. This is a method for extracting some topological information about the shape underlying a point cloud in Euclidean space at different scales. In Chapter 6, we give the basic topological background on homology, persistent homology, and related algorithms.

The new meshing methods developed in the first half of the thesis are particularly useful for geometric persistent homology. First, the mesh gives a good approximation to the types of smooth functions considered for geometric persistent homology, namely distance-like functions. Second, the algorithm may be run directly on the linear size hierarchical mesh, thus we can dispense with all of the output sensitive terms in the analysis, beating the Ruppert bound. This is important because it shows that for a real problem, relaxing the meshing quality definition still results in a useful complex. Chapter 7 covers the simple case when the desired function is the distance to the point cloud. Chapter 8 extends the mesh-based persistence method to more general distance functions considered in the literature, as well as to sequences of related distance functions.

Chapter 2

Basic Definitions and Notation

2.1 Points, Vectors, and Distances.

We will treat points of d-dimensional Euclidean space as vectors in $\mathbb{R}^d$. As such, we denote the Euclidean distance between two points $x, y \in \mathbb{R}^d$ as $|x-y|$. Moreover, we allow the usual operations of scalar multiplication and addition on points. For subsets of $\mathbb{R}^d$, the plus sign $(+)$ denotes the Minkowski sum, $A + B = \{a + b : a \in A, b \in B\}$. For example, if S is the unit sphere in $\mathbb{R}^d$ centered at the origin, c is any point, and r is a nonnegative real number, then $rS + \{c\}$ is the sphere of radius r centered at c. We will abuse notation slightly and write $rS + c$ when it is clear that c represents the singleton set $\{c\}$. We will also define the distance from a point x to a set S as $\mathbf{d}(x, S) = \inf_{y \in S} |x - y|$. The **diameter** of a closed set S is the maximum distance between any pair of points in S, $\mathbf{diameter}(S) = \max_{p,q \in P} |p - q|$. A set is **bounded** if $\mathbf{diameter}(S)$ is finite and **unbounded** otherwise. Let s_P denote the minimum distance between distinct points of P: $s_P = \min_{p,q \in P, p \neq q} |p - q|$. The ratio of the largest to smallest pairwise distances among point of P is called the **spread** and is denoted Δ. Formally,

$$\Delta = \frac{\mathbf{diameter}(P)}{s_P}.$$

A function $f : \mathbb{R}^d \to \mathbb{R}$ is **t-Lipschitz** if $f(x) \leq f(y) + t|x - y|$ for all $x, y \in \mathbb{R}^d$. More generally, for a map f between any pair of metric spaces $(X_1, \mathbf{d}_1)$ and $(X_2, \mathbf{d}_2)$, f is t-Lipschitz if $\mathbf{d}_2(f(x), f(y)) \leq t\mathbf{d}_1(x, y)$ for all $x, y \in X_1$. For example, the distance to a finite set of points $P \subset \mathbb{R}^d$ is 1-Lipschitz because $\mathbf{d}(x, P) \leq \mathbf{d}(y, P) + |x - y|$ for any $x, y \in \mathbb{R}^d$.

We write $\mathbf{ball}(c, r)$ to denote the open ball of radius r centered at c, and $\overline{\mathbf{ball}}(c, r)$ to denote the closed ball of radius r centered at c. The volume of a set X is denoted $\mathbf{Vol}(X)$. The volume of the d-dimensional unit ball is $\mathbb{V}_d := \mathbf{Vol}(\mathbf{ball}(0, 1))$. The **diameter** of a set X is the supremum of the distances between pairs of points of X and is denoted $\mathbf{diameter}(X)$.

Let $X \subset \mathbb{R}^d$ be any set and let $x = \sum_{x_i \in X} \lambda_i x_i$ be a linear combination of the points of X. Such a linear combination is **affine** if the sum of the coefficients is 1. It is **nonnegative** if the

coefficients are nonnegative. It is **convex** if it both affine and nonegative. The **affine closure** of X, denoted $\mathrm{aff}(X)$, is the set of all affine combinations of points of X. The **cone** of X, denoted $\mathrm{cone}(X)$, is the set of all nonnegative combinations of points in X. The **convex closure** of X, denoted $\mathrm{conv}(X)$, is the set of all convex combinations of points in X.

The intersection of a finite set of closed halfspaces is a **polyhedron**. The convex closure of a finite set of points is a **polytope**. The "Main Theorem" of polytopes says that a polytope is a polyhedron and a bounded polyhedron is a polytope [Zie95, Thm. 1.1]. The book by Ziegler [Zie95] is a good reference for polytope theory and we have adopted many of his notations and conventions in this thesis.

Let X be a polyhedron and let H be a halfspace containing X. If F is the hyperplane bounding H, then $Y = X \cap F$ is a **face** of X. The face Y is itself a polyhedron and its **dimension**, $\dim(Y)$ is the dimension of the affine space $\mathrm{aff}(Y)$. We also consider X to be a face of itself. The faces of dimension 0 are called **vertices**, faces of dimension 1 are called **edges** and the faces of dimension $\dim(X) - 1$ are **facets**. The empty set $\emptyset$ is face of every polyhedron and has dimension -1.

2.2 Complexes.

An **abstract simplicial complex** $\mathcal{K}$ is a family of subsets of a vertex set V that is closed under taking subsets. A subset $\sigma \in \mathcal{K}$ is called a **simplex** and its **dimension** is one less than its cardinality, $\dim(\sigma) = |\sigma| - 1$. The empty simplex $\{\emptyset\}$ is an element of every simplicial complex and its dimension is -1. A subset σ of a simplex σ' is itself a simplex and we say that σ is a **face** of σ'. The subset operation induces a partially ordered set (or **poset**) on the simplices in an abstract simplicial complex. Notice that an abstract simplicial complex is defined here without any reference to a geometric embedding.

When the vertex set of an abstract simplicial complex is a set of points in $\mathbb{R}^d$, there is a natural mapping of the complex to $\mathbb{R}^d$ by mapping to $\mathrm{conv}(\sigma)$. We get an **embedded simplicial complex** if for every pair of simplices (σ_1, σ_2),

$$\mathrm{conv}(\sigma_1 \cap \sigma_2) = \mathrm{conv}(\sigma_1) \cap \mathrm{conv}(\sigma).$$

The **underlying space** of an embedded simplicial complex $\mathcal{K}$ is $\bigcup_{\sigma \in \mathcal{K}} \mathrm{conv}(\sigma)$.

More generally, a **polyhedral complex** $\mathcal{C}$ is a finite collection of polyhedra such that

1. the empty polyhedron is in $\mathcal{C}$,
2. for all $X \in \mathcal{C}$, the faces of X are in $\mathcal{C}$, and
3. for all $X, Y \in \mathcal{C}$, $X \cap Y$ is a face of both X and Y.

2.3 Voronoi and Delaunay.

Let P be a finite set of points in $\mathbb{R}^d$. The **Voronoi cell** of $u \in P$ is the set of all points in $\mathbb{R}^d$ that have v as a nearest neighbor among the points of P. Formally, the Voronoi cell of u is defined as

$$\text{Vor}_P(u) = \{x \in \mathbb{R}^d : \mathbf{d}(x, P) = |x - u|\}.$$

Note that $\text{Vor}_P(u)$ is a closed, convex polyhedron, though it may be unbounded. The collection of Voronoi cells and their face posets decompose all of $\mathbb{R}^d$ into polyhedral complex called the **Voronoi diagram**, denoted Vor_P. We will often use the term **corner** rather than vertex to describe the 0-dimensional faces of the Voronoi cells to avoid ambiguity with the Delaunay vertices. In particular, every Voronoi cell $\text{Vor}_P(u)$ has a **farthest corner** which is the corner of the cell farthest from u.

The **Delaunay complex** is the dual complex to the Voronoi diagram. It has a cell for every maximal cospherical subset X of P such that the circumscribing ball of X contains no points of P in its interior. If no subset of $d + 2$ points of P are cospherical then we say that P is **in general position** and the resulting Delaunay complex will be an embedded simplicial complex known as the **Delaunay triangulation**. We will assume throughout that P is in general position. The circumball of a Delaunay simplex is the unique ball of minimum radius circumscribing the vertices. The circumballs of the Delaunay simplices are called **D-balls**. The centers of the D-balls are the corners of the Voronoi cells. The duality relationship between the Delaunay triangulation and the Voronoi diagram is useful for algorithms because it allows a single data structure to represent either (or both). The book by Edelsbrunner [Ede01] is a good resource for the basics of Delaunay triangulations, Voronoi diagrams, and related algorithms.

2.4 Mesh Generation.

Given a point $p \in P$ and its Voronoi cell $V = \text{Vor}_P(p)$, we define the **in-radius** r_p as the radius of the largest ball centered at p contained in V. Similarly, the **out-radius** R_p is the radius of the smallest ball centered at p that contains all of the corners of V. Thus, for unbounded Voronoi cells, the out-radius is still bounded. The **aspect ratio** of V is $\frac{R_p}{r_p}$. A set of points P is **τ-well-spaced** if for all $p \in P$, $\frac{R_p}{r_p} \leq \tau$.

The **meshing problem** for points is defined as follows.

- Input: a set of points P and a parameter τ.
- Output: a τ-well-spaced superset M of P and Vor_M.

Since the Voronoi diagram is determined uniquely by the point set, we will often refer to M as the mesh and let its induced Voronoi/Delaunay structure remain implicit. We adopt the convention that n denotes the input size, $|P|$, and m denotes the output size, $|M|$. In other literature on Voronoi diagrams, it is common to refer to the points of M as "sites", however, in this work, we adopt the intuition from mesh generation and call them **vertices** (the points of M are the vertices of the Delaunay triangulation).

The basic analytic tool used in mesh generation is the **local feature size** function, which assigns a nonnegative real number to every point in $\mathbb{R}^d$. It is defined with respect to a point set P as the distance to the second nearest neighbor:

$$\mathbf{f}_P(x) = \inf\{r : |\mathbf{ball}(x, r) \cap P| \geq 2\}.$$

In a good aspect ratio mesh M, the feature size at a vertex $v \in M$ gives an approximation to the in-radii of $\mathrm{Vor}_M(v)$. The feature size also induces a measure on $\mathbb{R}^d$ defined for any set $A \subset \mathbb{R}^d$ as

$$\int_A \frac{dx}{\mathbf{f}_P(x)^d}.$$

The integral above is known as the **feature size integral** and it is the primary tool in mesh size analysis.

It is often useful to have a weaker condition for point sets than τ-well-spaced. One such condition is provided by the notions of medial points and well-paced orderings defined as follows. Let P be a set of points and let $P' = P \cup \{q\}$ for some $q \notin P$. The point q is **θ-medial** with respect to P if $\frac{\mathbf{f}_{P'}(q)}{\mathbf{f}_P(q)} \geq \theta$. Equivalently, the ratio of the nearest to second nearest neighbors to q in P is at least θ. An ordering on a set $P \subset \mathbb{R}^d$ is a **θ-well-paced** ordering if

1. $|p_1 - p_2| = \mathbf{diameter}(P)$, and
2. p_i is θ-medial with respect to $\{p_1, \ldots, p_{i-1}\}$ for all $i = 1 \ldots n$.

The term well-paced refers to the rate of change of the feature size function as points are added according to the given ordering. In this case, no single point causes a drastic change in the feature size. We will prove in Chapter 3 that well-paced points in a reasonably sized bounding domain can be augmented with only a constant factor more points to produce a well-spaced mesh.

We augment the input points P with points on the surface of a **bounding box** or a **bounding ball** of extra points to ensure that the boundary of the point set is uniform. This is a standard practice in mesh generation, and in previous work, we showed that the effect on the total mesh size is negligible [HMPS09]. We will always choose the diameter of the bounding ball to be no more than a constant times larger than the diameter of P.

2.5 Notation

Here is a reference list of common notation used throughout the text.
d: the ambient dimension of the input
$\mathbb{R}$: the real numbers
$\mathbb{R}_{\geq 0}$: the nonnegative real numbers
$\mathbb{R}^d$: the affine space of d-dimensional points
P: a set of points in $\mathbb{R}^d$
$\mathbf{ball}(c, r)$: the open ball of radius r centered at c
$\overline{\mathbf{ball}}(c, r)$: the closed ball of radius r centered at c
$\mathbf{diameter}(S)$: the distance between the farthest pair of points of S
$|S|$: the cardinality of the set S
n: the size of the input set P, i.e. $n = |P|$
M: a superset of P
Ω: The bounding domain, a subset of $\mathbb{R}^d$
Vor_S: The Voronoi diagram of a point set S
$\text{Vor}_S(p)$: The Voronoi cell of the point p in $\text{Vor}(S)$.
$\text{Vor}(p)$: The Voronoi cell of the point p in $\text{Vor}(S)$, when S is assumed
Del_S: The Delaunay triangulation of S
$\mathbf{Vol}(X)$: The volume of the set X
$\mathbb{V}_d$: The volume of the d-dimensional unit ball
$\mathcal{K}$: A simplicial complex
$\partial(X)$: The boundary of the set X

Chapter 3

Mesh Size Analysis

3.1 Introduction

Many of the results in this thesis depend on our ability to bound (both from above and below) the number of vertices in the output of the Voronoi Refinement algorithm. In its simplest form, this algorithm starts with the Voronoi diagram of the input and repeatedly adds the farthest corner of any Voronoi cell with aspect ratio greater than some constant $\tau > 2$ until all have good aspect ratio. The remarkable simplicity of this algorithm has led to its widespread use[1], but its analysis is nontrivial. In this chapter, we present a new way to compute asymptotically tight per-instance bounds on the output size.

We begin in Section 3.2 with a description of the Voronoi refinement algorithm. Then, in Section 3.3, we prove bounds on the feature size of the output of htis algorithm. This is the first step towards a generalization of the classical mesh sizing bounds of Ruppert [Rup95] in Section 3.4. Ruppert's analysis was limited to two-dimensions. We give extend it to higher dimensions and give special attention to bounding the dependence on the dimension. Such an analysis surely exists in the folklore, but we present it for three important reasons. One, we need a clear statement and proof from which to build new theorems. Two, it is the starting point and gives some intuition for our instance-optimal bounds. Three, it gives a simple instance of the pattern that we will follow for our later analysis of a new, state-of-the-art meshing algorithm.

To get Voronoi refinement to terminate, one must be a bit careful around the boundary. In Section 3.5, we

Well-spaced points require that *all* of the near neighbors of every point are approximately the same distance away. For well-paced points, we only require that the first two nearest neighbors are approximately the same distance away. In Section 3.8, we show how to bound the feature size integral by looking at how medial each input vertex is. This allows us to derive the desired per-instance bounds.

[1]This algorithm is most commonly seen in its dual formulation as Delaunay refinement.

3.2 The Basic Voronoi Refinement Algorithm

Voronoi refinement has its roots in a closely related algorithm known as **Delaunay refinement**, first developed by Paul Chew for 2-dimensional meshing [Che89]. In Delaunay refinement, the algorithm iteratively adds Steiner points to a Delaunay triangulation as long as there are triangles with small angles. To eliminate a triangle with a small angle, a Steiner point is added at the circumcenter of the vertices of the triangle. This process continues until all angles are larger than some user-defined constant. In dimensions higher than 2, the small angle condition is replaced with the radius-edge condition [MTTW99].

The dual view of Delaunay refinement is Voronoi refinement. As long as there is a Voronoi cell of aspect ratio greater than some constant, the farthest corner of the offending Voronoi cell is added.

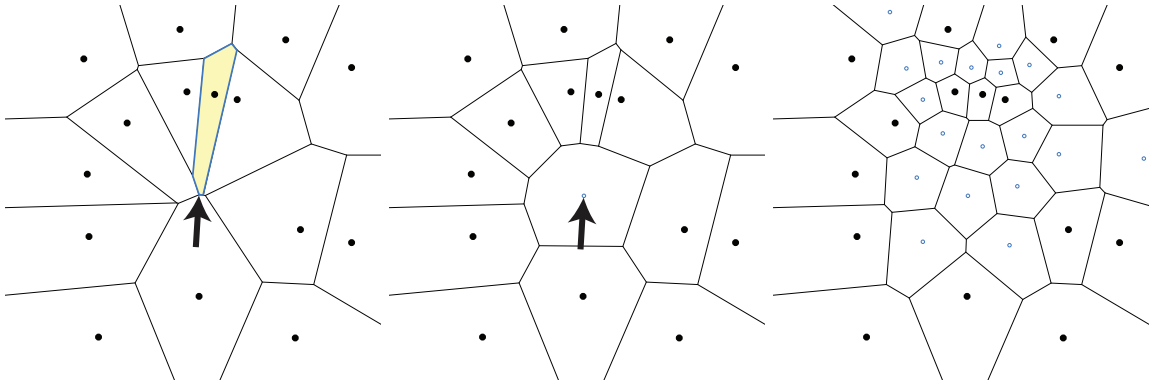

Figure 3.1: In the VoronoiRefine algorithm, a Voronoi cell of bad aspect ratio is identified (left). The farthest corner of the offending cell is added to the Voronoi diagram (center). This process is repeated until all of he Voronoi cell have good aspect ratio (right).

Recall that the Voronoi/Delaunay duality implies that Voronoi cell corners are exactly the circumcenters of Delaunay simplices. The algorithms are not identical, but a Voronoi aspect ratio of τ implies a radius-edgs ratio bound of 2τ.

The algorithm takes an optional set of extra points N that are permitted to have aspect ratio τ or greater. Usually, N provides an outer boundary to the input set. In the course of refinement, the algorithm ignores the vertices of N. We give an explicit construction for a choice of N in Section 3.5 that guarantees termination. However, for many critical pieces of the analysis, the choice of N is arbitrary.

The basic Voronoi refinement algorithm is as follows. It is illustrated in Figure 3.1

VoronoiRefine (P, N, τ)

let $M \leftarrow P \cup N$

while there is some $v \in M \setminus N$ with aspect ratio of $\text{Vor}(v) > \tau$

let c be the farthest corner of $\text{Vor}(v)$.

let $M \leftarrow M \cup \{c\}$
return M

Note that the algorithm does not specify the order in which in the bad aspect ratio cells are processed. One popular choice is to refine the largest cells first. This is sometimes called *top-down refinement*.

3.3 Bounding the feature size of the output

It is not obvious that Voronoi refinement terminates, because for some inputs and choices of τ, it will run forever. In this section, we prove that the feature size function $\mathbf{f}_M$, defined with respect to the output set $M = \textsc{VoronoiRefine}(P, N, \tau)$ is not too small compared to the input feature size $\mathbf{f}_P$. This is the first step towards showing that the algorithm terminates and moreover in Section 3.4 that the output has optimal size.

For an ordered set of points M with order relation $\prec$, the **insertion radius** λ_v of $v \in M$ is the distance to its nearest predecessor:

$$\lambda_v = \min_{u \prec v} |u - v|.$$

The term insertion radius comes from the algorithm; it is the largest radius r such that $\mathbf{ball}(v, r)$ contains no previously inserted vertices at the time v is inserted.

The following lemma shows it suffices to bound the insertion radii of a set of points in order to bound the output feature size.

Lemma 3.3.1. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a t-Lipschitz function and let $K > 0$ be a constant, Let M be an ordered set of points and let $N \subset M$ be a prefix of the ordering. If $f(v) \leq (K - t)\lambda_v$ for all $v \in M \setminus N$ then*

$$f(v) \leq K\mathbf{f}_M(v)$$

for all $v \in M \setminus N$.

Proof. Let $v \in M \setminus N$ be any point and let u be its nearest neighbor in M (other than itself) so that $\mathbf{f}_M(v) = |u - v|$. Let $\prec$ denote the order relation on M. If $u \prec v$ then $f(v) \leq K\lambda_v = K\mathbf{f}_M(v)$. So, we may assume $v \prec u$, and so $\lambda_u \leq |u - v|$ and $u \notin N$. So, $f(v) \leq f(u) + t|u - v| \leq (K - t)\lambda_u + t|u - v| \leq K\mathbf{f}_M(v)$. □

To apply Lemma 3.3.1 in the analysis of VoronoiRefine, the natural choice is $f = \mathbf{f}_P$ which is 1-Lipschitz. This allows us to prove the main theorem of this section.

Theorem 3.3.2. *Let $P \subset \mathbb{R}^d$ be a set of at least two points and let N be an optional set of boundary points such that $\mathbf{f}_P(v) = \mathbf{f}_{P \cup N}(v)$ for all $v \in P$. Let $M = \textsc{VoronoiRefine}(P, N, \tau)$ for $\tau > 2$. For all $v \in M \setminus N$,*

$$\mathbf{f}_P(v) \leq K\mathbf{f}_M(v),$$

where $K = \frac{2\tau}{\tau - 2}$.

Proof. Lemma 3.3.1 implies that it suffices to show that $\mathbf{f}_P(u) \leq (K-1)\lambda_u$ for all $u \in M \setminus N$. We will prove this by induction on the number of Steiner points added. Let M_i denote the mesh vertices after i Steiner points have been added. In the base case, $M_0 \setminus N = P$, so $\mathbf{f}_P(u) = \mathbf{f}_{P \cup N}(u) \leq \lambda_u$ for all $u \in M_0 \setminus N$, which is less than $(K-1)\lambda_u$ because $K > 2$.

Assume inductively that $\mathbf{f}_P(u) \leq (K-1)\lambda_u$ for all $u \in M_i \setminus N$. Let v be $(i+1)$st Steiner point added. Let $x \in M_i \setminus N$ be a vertex whose Voronoi cell had poor quality, initiating the insertion of v. Let y be the nearest neighbor of x in M_i.

$$\begin{aligned}
\mathbf{f}_P(v) &\leq \mathbf{f}_P(x) + |x - v| && [\mathbf{f}_P \text{ is 1-Lipschitz}] \\
&\leq K\mathbf{f}_{M_i}(x) + |x - v| && [\text{by induction and Lemma 3.3.1}] \\
&\leq K|x - y| + |x - v| && [\text{definition of } y] \\
&\leq \left(\frac{2K}{\tau} + 1\right)|x - v| && [\text{Vor}(x) \text{ had aspect ratio} > \tau] \\
&= (K-1)|x - v| && \left[K = \frac{2\tau}{\tau - 2}\right] \\
&= (K-1)\lambda_v. && [\text{definition of } x]
\end{aligned}$$

□

The proof of Theorem 3.4.1 reveals an important property of Delaunay/Voronoi refinement: there is nothing sacrosanct about corners. The proof only requires that $|x - y| \leq \frac{2}{\tau}|u - v|$. Thus, the algorithm can be modified to add any point of $\text{Vor}(v)$ that is sufficiently far from v and the bounds will still apply. This property is exploited in many algorithms to achieve meshes that are smaller [Üng09], conform to higher order features [She97, She02a, Rup95], and can be computed quickly [HMP06].

This theorem is almost sufficient to imply the termination of the VoronoiRefine algorithm. If we also have the guarantee that all Steiner points are contained in some compact subset $\Omega \subset \mathbb{R}^d$, then termination is guaranteed by a packing argument; each pair of output points has some minimum separation and the total volume of Ω is bounded. This guarantee is proven in the next section.

3.4 Optimality of the Voronoi Refinement Algorithm

Throughout, P will denote the set of input points in $\mathbb{R}^d$ and $M = \text{VoronoiRefine}(P, N, \tau)$ will denote the output vertices for some $\tau > 2$ independent of the dimension. The goal of this section is to present upper and lower bounds on $|M|$. The upper bounds are achieved by the VoronoiRefine algorithm. The lower bounds hold for all τ-well-spaced supersets of P subject to mild boundary conditions. Up to constant factors, the upper and lower bounds match.

3.4.1 Upper Bounds

We prove the upper bounds in a bit more generality. Rather than sticking to a particular sizing function such as the local feature size, we show that the same arguments will work for any Lipschitz function that bounds the size of the Voronoi cells. At first this may seem strange because the bounds are first given in a form completely independent of the input set P. Then we will show that plugging $\mathbf{f}_P$, the feature size with respect to the input, into the bounds gives the desired result. The more general conditions will be useful later.

Theorem 3.4.1. *Let P be a set of points in $\mathbb{R}^d$ and let $\Omega \subset \mathbb{R}^d$ be a convex set. Let $M \supset P$ be a set of mesh vertices and let $M' \subset M$ be the subset of vertices v such that $\mathrm{Vor}_M(v) \subseteq \Omega$. Let $f : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ be a 1-Lipschitz function. If $f(v) < c\mathbf{f}_M(v)$ for some constant c and all vertices $v \in M'$, then*

$$|M'| < \left(\frac{(2c+1)^d}{\mathbb{V}_d}\right) \int_\Omega \frac{dx}{f(x)^d}.$$

Proof. Recall that $r_v = \frac{\mathbf{f}_M(v)}{2}$ denotes the in-radius of $\mathrm{Vor}_M(v)$. Since f is 1-Lipschitz, the hypothesis that $f(v) < c\mathbf{f}_M(v)$ and the definition of $\mathbf{f}_M$ imply

$$f(x) < (2c+1)r_v \tag{3.1}$$

for all $v \in M'$ and all $x \in \mathbf{ball}(v, r_v)$. We can now derive the following upper bound.

$$\begin{aligned}
\int_\Omega \frac{dx}{f(x)^d} &\geq \sum_{v \in M'} \int_{\mathrm{Vor}_M(v)} \frac{dx}{f(x)^d} && \left[\mathrm{Vor}_M(v) \subseteq \Omega \text{ for all } v \in M'\right] \\
&\geq \sum_{v \in M'} \int_{\mathbf{ball}(v,r_v)} \frac{dx}{f(x)^d} && [f \text{ is nonnegative and } \mathbf{ball}(v, r_v) \subset \mathrm{Vor}(v)] \\
&> \sum_{v \in M'} \int_{\mathbf{ball}(v,r_v)} \frac{dx}{((2c+1)r_v)^d} && [\text{by (3.1)}] \\
&= \sum_{v \in M'} \frac{\mathbb{V}_d r_v^d}{(2c+1)^d r_v^d} && \left[\int_{\mathbf{ball}(v,r_v)} dx = \mathbb{V}_d r_v^d\right] \\
&= |M'| \left(\frac{\mathbb{V}_d}{(2c+1)^d}\right) && [\text{simplify}]
\end{aligned}$$

□

3.4.2 Lower Bounds

As with the upper bounds, the lower bounds are with respect to any convex subset $\Omega \subset \mathbb{R}^d$. We prove that if there is a subset of vertices whose Voronoi cells have bounded aspect ratio and cover Ω then there must be some minimum number of them which depends on the aspect ratio bound and the feature size integral.

Theorem 3.4.2. *Let P be a set of points in $\mathbb{R}^d$ and let $\Omega \subset \mathbb{R}^d$ be a convex set. Let $M \supset P$ be a set of mesh vertices and let $M' \subset M$ be the subset of vertices v such that $\mathrm{Vor}_M(v) \cap \Omega \neq \emptyset$. If $\mathrm{Vor}_M(v)$ is bounded and has aspect ratio at most τ for all $v \in M'$, then*

$$|M'| > \left(\frac{1}{\tau^d \mathbb{V}_d}\right) \int_\Omega \frac{dx}{\mathbf{f}_P(x)^d}.$$

Proof. For all $v \in M'$ and all $x \in \mathrm{Vor}(v)$,

$$\mathbf{f}_P(x) \geq \mathbf{f}_M(x) \geq r_v. \tag{3.2}$$

Because the Voronoi cells of M' are τ-quality, for each $v \in M'$,

$$\mathrm{Vor}(v) \subset \mathbf{ball}(v, \tau r_v). \tag{3.3}$$

$$\begin{aligned}
\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} &\leq \sum_{v \in M'} \int_{\mathrm{Vor}(v)} \frac{dx}{\mathbf{f}_P(x)^d} && \text{[Cover } \Omega \text{ with Voronoi cells]} \\
&\leq \sum_{v \in M'} \int_{\mathrm{Vor}(v)} \frac{dx}{r_v^d} && \text{[by (3.2)]} \\
&< \sum_{v \in M'} \int_{\mathbf{ball}(v, \tau r_v)} \frac{dx}{r_v^d} && \text{[by (3.3)]} \\
&= \sum_{v \in M'} \frac{\mathbb{V}_d \tau^d r_v^d}{r_v^d} && \text{[evaluate the integral]} \\
&= |M'| \left(\tau^d \mathbb{V}_d\right) && \text{[simplify]}
\end{aligned}$$

□

3.5 The bounding domain and its boundary net

The Voronoi diagram decomposes all of $\mathbb{R}^d$, but we want to restrict our attention only to a bounded subset. Moreover, it may continue to refine Voronoi cells everywhere in the space. The upper bounds of Section 3.4 show that the algorithm has a kind of "local termination". For any bounded subset of space, there will only be a constant number of mesh vertices placed there.

There are many approaches to dealing with the possibly unbounded behavior of mesh refinement algorithms. For example, some authors [CDE+00] have considered periodic point sets where every point x in the box $[0,1)^d$ is treated as an equivalence class of points $x + z$ for $z \in \mathbb{Z}^d$. Another popular approach is to contain the points in a bounding domain such as a box or a ball. For the case of bounding balls, Cohen-Steiner et al. [CSdVY04] introduced the *Split-on-sphere* method which causes new points near the boundary to get placed exactly on the boundary instead. In that context

the goal was not to curb infinite behavior on the outside of the mesh but rather to deal with infinite refinement near small angles in the input for meshes that conform not just to point sets but also to higher order features. This is the same problem turned inside out and numerous methods have been proposed to produce a finite mesh in these settings [CP03, PW04, CDR05, CDRR05, RW08, Phi09].

Other simple approaches can also be employed to keep the mesh refinement bounded. For example, it suffices to choose a sufficiently large ball containing the input and refuse to add any Steiner points outside that ball. The local termination upper bounds of Section 3.4 guarantee that only a finite number of points will be added and thus the algorithm will terminate.

In this section, we give a specific construction of a subset $\Omega \subset \mathbb{R}^d$ and a collection of points $N \subset \partial\Omega$ that surround the input. We then show that the vertices of VoronoiRefine(P, N, τ) stay bounded within Ω throughout the algorithm. This is based on a idea first proposed by Hudson [Hud07], though we will get slightly tighter bounds, partly because the VoronoiRefine algorithm as stated in this chapter, does not attempt to refine these boundary vertices. Later, in Chapter 4, we will use a similar set of bounding points in order to enclose small subsets of the input. There, we will find that these subsets are critical to producing linear size meshes for a slightly relaxed notion of quality.

For the input set P, we assume without loss of generality that some $p \in P$ is the origin. Let

$$\delta := \textbf{diameter}(P).$$

Let

$$r_\Omega := (K\sqrt{2} + 2)\delta,$$

where $K = \frac{2\tau}{\tau - 2}$ and τ is the desired Voronoi aspect ratio bound. The **bounding domain** of the set P is

$$\Omega := \textbf{ball}(0, r_\Omega).$$

Let $\overline{\Omega} := \Omega \cup \partial\Omega$ denote the closure of Ω. The **outer annulus** of Ω is

$$A_\Omega := \Omega \setminus \textbf{ball}(0, r_\Omega - \delta).$$

The bounding domain and outer annulus are illustrated in Figure 3.2.

We must be careful to keep the refinement process bounded within Ω. Some Voronoi cells are unbounded, however, it will suffice to guarantee that this set is small and moreover, that all of the corners of the Voronoi diagram are contained in Ω. We will explicitly pick a small set of vertices on the boundary of Ω and then prove that our choice of bounding points keeps the refinement process contained in the domain.

The **boundary net** of Ω is a maximal set of points $N \subset \partial\Omega$ such that

1. for all $x \in \partial\Omega$, $\mathbf{d}(x, N) \leq \delta$, and
2. for all $u, v \in N$, $|u - v| > \delta$.

The first condition is a covering constraint; it says that the balls of radius δ centered at points

of N cover $\partial\Omega$. The second condition is a packing condition; it says that the balls of radius $\frac{\delta}{2}$ centered at points of N are disjoint. Such a set N is sometimes referred to as **metric δ-net**. It can be constructed by a simple greedy algorithm, at each step adding a point of $\partial\Omega$ farthest from the current set. Up to constant factors this algorithm is optimal and gives a net of size $|N| \leq K^{O(d)}$ (see Matousek [Mat02][Lemma 13.1.1] for a proof). In fact, this greedy algorithm is a kind of Voronoi refinement restricted to the sphere because the farthest point from the current set is a corner of the Voronoi diagram on the sphere.

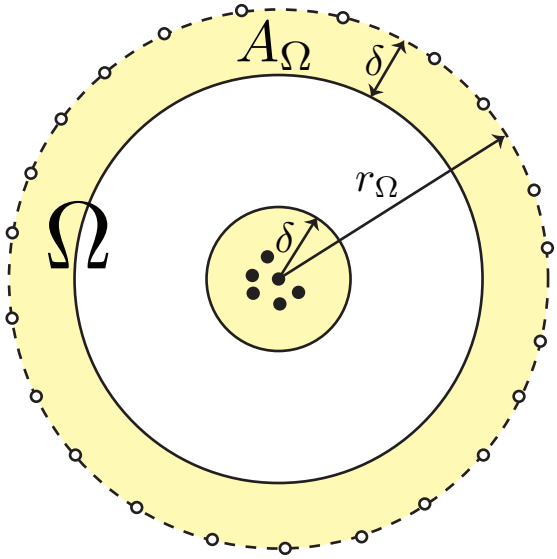

Figure 3.2: The black points are the input points P and the white points on the boundary of Ω are the boundary net N.

For a point set M, there is no general bound on the diameter of the set of corners of Vor_M as a function only of $\mathbf{diameter}(M)$. However, choosing a boundary net in this way guarantees that all of the corners of the Voronoi diagram stay close in the course of the VoronoiRefine algorithm. Towards proving this guarantee, we start with the following lemma.

Lemma 3.5.1. *Let Ω be the bounding domain of a point set P and let N be a boundary net of Ω. If M is set of points such that $P \cup N \subset M \subseteq \overline{\Omega}$ and no vertex of M lies in the outer annulus A_Ω then all corners of Vor_M are contained in Ω.*

Proof. Suppose for contradiction that $c \notin \Omega$ for some corner c of Vor_M as in Figure 3.3. The corner c has a corresponding Delaunay simplex $\sigma \subset M$ containing at least one vertex $v \in M \setminus N$. Let z be the intersection of the segment $\overline{cv}$ with $\partial\Omega$. Observe that $\mathbf{ball}(c, |c-v|)$ is a Delaunay ball so it is empty of points of M. Since $\mathbf{ball}(z, |z-v|) \subset \mathbf{ball}(c, |c-v|)$, it is also empty of points of M and so $|z-v| = \mathbf{d}(z, M)$. The covering condition on the boundary net guarantees that $\mathbf{d}(z, N) \leq \delta$. So, we conclude

$$\mathbf{d}(v, \partial\Omega) \leq |z-v| = \mathbf{d}(z, M) \leq \mathbf{d}(z, N) \leq \delta,$$

which implies $v \in A_\Omega$, a contradiction. □

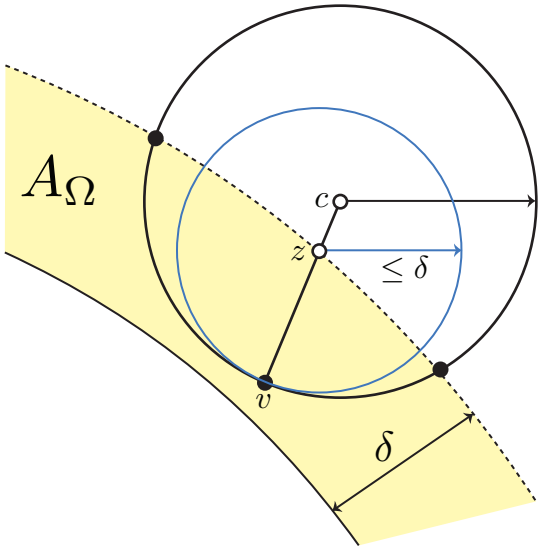

Figure 3.3: A Delaunay ball with circumcenter c that lies outside of Ω has a vertex $v \in M \setminus N$ on its boundary that lies in the outer annulus A_Ω.

The preceding lemma implies that as long as no vertices are added to the outer annulus in the course of the algorithm, then no corner will lie outside of Ω. This fact allows us to prove the following Lemma, which guarantees that the output of VoronoiRefine(P, N, τ) is contained in $\overline{\Omega}$.

Lemma 3.5.2. *If $M =$ VoronoiRefine(P, N, τ) for a domain Ω then $M \subset \overline{\Omega}$.*

Proof. The proof will be by induction on the number of Steiner points. Assume inductively that the mesh vertices M_i after i Steiner points have been added are contained in $\overline{\Omega} \setminus A_\Omega$. The base case is a mesh composed of only the points $P \cup N$, which which satisfy the inductive hypothesis by construction. Let v be the $(i+1)$st Steiner point added. By induction and Lemma 3.5.1, we know that $v \in \Omega$. Suppose for contradiction that $v \in A_\Omega$. Since v is a Voronoi corner, it has $d+1$ nearest neighbors at the time it is inserted at least one of which is not in N; call it u. We first bound $|u - v|$ from below using the bounds on the feature size.

$$\begin{aligned} |u - v| &\geq \mathbf{f}_M(v) && [u, v \in M] \\ &\geq \frac{\mathbf{f}_P(v)}{K} && [\text{by Theorem 3.3.2}] \\ &> \frac{r_\Omega - 2\delta}{K}. && [\mathbf{f}_P(x) > (r_\Omega - 2\delta) \text{ for all } x \in A_\Omega] \end{aligned}$$

We now bound $|u - v|$ from above using the conditions on the boundary net.

$$\begin{aligned} |u - v| &\leq \mathbf{d}(v, N) && [N \subset M_i] \\ &\leq \delta\sqrt{2}. && \left[\mathbf{d}(x, N) \leq \delta\sqrt{2} \text{ for all } x \in A_\Omega\right] \end{aligned}$$

The preceding inequalities imply that

$$\frac{r_\Omega - 2\delta}{K} < \delta\sqrt{2},$$

and thus, $r_\Omega < \delta(K\sqrt{2}+2) = r_\Omega$, a contradiction. □

3.6 Termination of Voronoi Refinement

Theorem 3.3.2 shows that the feature size of the output of Voronoi refinement is bounded from below by a constant times the feature size of the input. The upper bound of Theorem 3.4.1 says that any set of points satisfying such a bound in a convex set Ω, has size bounded from above by a constant times the feature size integral. Moreover, the lower bound of Theorem 3.4.2 says that the number of vertices in Ω for any bounded aspect ratio Voronoi diagram must be at least a constant times the feature size integral. By choosing the domain Ω and the boundary net N as in Section 3.5, we can contain all of the mesh vertices to lie within Ω. Putting these together, we get upper and lower bounds on the output size of VoronoiRefine(P, N, τ) in the following theorem.

Theorem 3.6.1. *Let $P \subset \mathbb{R}^d$ be a finite set of points, let $\Omega \subset \mathbb{R}^d$ and $N \subset \partial\Omega$ be a boundary domain and net as constructed in Section 3.5, and let τ be a constant greater than 2. If $M =$* VoronoiRefine(P, N, τ) *then*

$$C\tau^{-d} < |M| < C(2K+1)^d + |N|,$$

where $C = \frac{1}{\mathbb{V}_d}\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d}$ and $K = \frac{2\tau}{\tau-2}$.

Proof. The upper bound is a direct application of Theorem 3.4.1 using $f = \mathbf{f}_P$.

For the lower bound, we break the feature size integral into two parts by partitioning Ω into $B = \mathbf{ball}(0, \frac{r_\Omega}{2})$ and $\Omega \setminus B$. Voronoi cells from the boundary net N do not intersect B and all other Voronoi cells have aspect ratio at most τ. So, Theorem 3.4.2 implies that

$$|M \setminus N| > \left(\frac{1}{\tau^d \mathbb{V}_d}\right) \int_B \frac{dx}{\mathbf{f}_P(x)^d}. \tag{3.4}$$

Moreover, $\mathbf{f}_P(x) \geq \frac{r_\Omega}{3}$ for all $x \in \Omega \setminus B$ (slightly better constants are possible, but these will

suffice). So,

$$\begin{aligned}
\int_{\Omega\setminus B} \frac{dx}{\mathbf{f}_P(x)^d} &\leq \int_{\Omega\setminus B} \frac{3^d dx}{r_\Omega^d} && \left[\mathbf{f}_P(x) \geq \frac{r_\Omega}{3}\right] \\
&\leq d\mathbb{V}_d \left(\frac{3}{r_\Omega}\right)^d \int_{\frac{r_\Omega}{2}}^{r_\Omega} r^{d-1} dr && \text{[rewrite in polar coordinates]} \\
&= \left(3^d - \left(\frac{3}{2}\right)^d\right) \mathbb{V}_d. && \text{[evaluate]} \\
&< |N|\mathbb{V}_d. && \left[|N| > 3^d\right] \qquad (3.5)
\end{aligned}$$

To complete the proof, we sum the bounds on the integral over B and $\Omega \setminus B$:

$$\begin{aligned}
\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} &= \int_B \frac{dx}{\mathbf{f}_P(x)^d} + \int_{\Omega\setminus B} \frac{dx}{\mathbf{f}_P(x)^d} && \text{[partition } \Omega\text{]} \\
&< |M \setminus N|(\tau^d \mathbb{V}_d) + |N|\mathbb{V}_d && \text{[by (3.4) and (3.5)]} \\
&< |M|(\tau^d \mathbb{V}_d).
\end{aligned}$$

□

The two main ingredients in this bound will be seen again in Chapter 4: a bound on the output feature size in terms of the input feature size and the integral of the input feature size over the domain. One interpretation is that the feature size induces a measure on the space which approximately counts the number of mesh vertices. The total measure is the feature size integral and thus the total number of vertices is bounded by computing this integral. In Section 3.8, we give a novel method for bounding this integral, and thus for bounding the size of an optimal mesh.

3.7 Over-Refinement

The mesh sizing bounds of Section 3.4 apply for a very broad class of functions, not just the local feature size function. Often, it is useful to refine the mesh to a scale much smaller than the local feature size, in particular for the function $f_\varepsilon(x) = \varepsilon \mathbf{f}_P(x)$. Theorem 3.4.1 applies also in this case. Since this will come up repeatedly, it is worthwhile to make the analysis explicit and to explain the associated algorithm.

According to Theorem 3.4.1, we know that

$$|M_\varepsilon| = \mathbb{V}_d O\left(\int_\Omega \frac{dx}{f_\varepsilon(x)^d}\right),$$

for any mesh M_ε sized according to f_ε. Using the definition of f_ε, this implies that

$$|M_\varepsilon| = \varepsilon^{-d} O(|M|),$$

where M is sized according to the local feature size.

To compute M_ε, it suffices to be able to compute f_ε, or equivalently $\mathbf{f}_P$. Then, the construction is straightforward:

1. Construct M.
2. Refine the Voronoi cell of a vertex v if $r_v > f_\varepsilon(v)$. Otherwise, terminate.
3. Add Steiner points until quality is achieved.
4. Return to step 2.

To show that this algorithm terminates with the desired bounds, we must prove a generalization of Theorem 3.3.2, which applies to any Lipschitz function $f \leq K\mathbf{f}_P$. The upper bound on f is just the bound achieved when no extra refinement happens.

Theorem 3.7.1. *Let $f \leq (K-t)\mathbf{f}_P$ be a t-Lipschitz function, where $K = \max\{\frac{2t\tau}{\tau-2}, 1+2t\}$. Let M be a mesh constructed from P by Voronoi Refinement with over-refinement down to size f with bounding net N. For all $v \in M \setminus N$,*

$$f(v) \leq K\mathbf{f}_M(v).$$

Proof. Lemma 3.3.1 implies that it suffices to show that $f(v) \leq (K-t)\lambda_v$ for all $v \in M \setminus N$. We will prove this by induction on the number of Steiner points. Let M_i be the mesh after i Steiner points have been added. In the base case, $M_0 \setminus N = P$, so $f(v) \leq (K-t)\mathbf{f}_P(v) \leq (K-t)\lambda_v$ for $v \in M_0 \setminus N$.

Assume inductively that $f(u) \leq (K-t)\lambda_u$ for all $u \in M_i$. Let v be $(i+1)$st Steiner point as the farthest corner of $\mathrm{Vor}(x)$ for some $x \in M_i \setminus N$. There are two cases to consider, either $\mathrm{Vor}_{M_i}(x)$ had aspect ratio greater than τ or $\mathrm{Vor}(x)$ was too large compared to $f(x)$. The former case, as we will see, is nearly identical to the main step in Theorem 3.3.2.

Case 1: $\mathrm{Vor}_{M_i}(x)$ had bad aspect ratio. Let y be the nearest neighbor of x in M_i.

$$\begin{aligned}
f(v) &\leq f(x) + t|x-v| && [f \text{ is } t\text{-Lipschitz}] \\
&\leq K\mathbf{f}_{M_i}(x) + t|x-v| && [\text{by induction and Lemma 3.3.1}] \\
&\leq K|x-y| + t|x-v| && [\text{definition of } y] \\
&\leq \left(\frac{2K}{\tau} + t\right)|x-v| && \left[|x-y| < \frac{2|x-v|}{\tau} \text{ by our algorithm}\right] \\
&\leq (K-t)|x-v| && \left[K \geq \frac{2t\tau}{\tau-2} \text{ by definition}\right] \\
&= (K-t)\lambda_v. && [\text{by the definition of } x]
\end{aligned}$$

Case 2: $\mathrm{Vor}_{M_i}(x)$ was too large compared to $f(x)$. Too large means that $r_x > f(x)$ where r_x is the in-radius of $\mathrm{Vor}(x)$. Now, $|x-v| \geq r_x$, because v is on the boundary of $\mathrm{Vor}_{M_i}(x)$ and $\mathbf{ball}(x, r_x) \subset \mathrm{Vor}_{M_i}(x)$. So, $f(x) \leq |x-v|$, allowing us to derive the following to complete the

proof.

$$\begin{aligned} f(v) &\leq f(x) + t|x - v| && [f \text{ is 1-Lipschitz}] \\ &\leq (1+t)|x - v| && [f(x) \leq |x - v|] \\ &\leq (1+t)\lambda_v && [\text{by the definition of } x] \\ &\leq (K - t)\lambda_v. && [K \geq 1 + 2t \text{ by definition}] \end{aligned}$$

□

The preceding Theorem, implies a general method for over-refining meshes down to some given sizing function f. The primary limitation to using many such functions is our ability to compute them efficiently. For the special case of f_ε, it suffices to compute $\mathbf{f}_P$. After constructing M, $\mathbf{f}_P(x)$ can be computed in constant time by simply identifying Voronoi cell of Vor_M containing x. Moreover, this only takes constant time, because each insertion after M is constructed is the corner of a Voronoi cell of a vertex v and the nearest neighbor of v in M has already been computed. Both of these facts depend on the constant complexity of cells in good aspect ratio Voronoi diagrams (see for example [MTTW99, HMP06] or in Chapter 4 for a more general setting). Thus, the total time to do the over-refinement is $O(|M_\varepsilon|)$ after M has been computed.

3.8 Analysis for well-paced points

From the analysis in Section 3.4, the mesh size is within a constant factor of the feature size integral. It's not hard to check that the feature size integral over all of $\mathbb{R}^d$ goes to to infinity, because if the domain Ω to be meshed is arbitrarily large, then the output must also be arbitrarily large. Consequently, we restrict our attention to bounding domains Ω that are only a constant factor larger than the diameter of the input. Even in such cases, the feature size integral is not necessarily bounded by any function of n. It is not too difficult to prove that the integral is always bounded from above by $n \log \Delta$ where Δ is the spread of the input, but that analysis is rarely tight. Instead, we prove tight, per-instance bounds on the mesh size.

Our analysis has the added advantage that it makes clear what regions of the domain cause the mesh size to go superlinear. Thus, it gives us a clear guide on how to relax the usual mesh quality conditions to guarantee a linear size mesh, as we will see in Chapter 4. The main result of this section, Theorem 3.8.1, gives us a way to compute the mesh size. These bounds are tight. The main tools of the analysis are medial points and well-paced orderings (see page 10 for definitions).

We can now state the main result of this section. Its proof will be broken up into two parts: the upper bound in Lemma 3.8.5 and the lower bound in Lemma 3.8.7.

Theorem 3.8.1. *Let $P = \{p_1 \ldots, p_n\}$ be an ordered set of points such that $|p_1 - p_2| = \textbf{diameter}(P)$ and let $P_i = \{p_1, \ldots, p_i\}$ be the prefixes of this ordering for $i = 1 \ldots n$. Let $\Omega \subset \mathbb{R}^d$ be a bounding*

domain such that

$$\mathbf{ball}(p_1, 3\mathbf{diameter}(P)) \subseteq \Omega \subseteq \mathbf{ball}(p_1, c\mathbf{diameter}(P)),$$

for some constant $c \geq 3$.

$$\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} = \Theta\left(\sum_{i=3}^{n} \ln \frac{1}{\theta_i}\right) - O(n),$$

where $\theta_i = \frac{\mathbf{f}_{P_i}(p_i)}{\mathbf{f}_{P_{i-1}}(p_i)}$.

Up to the additive factor of n, the bound is tight. This theorem and Theorem 3.6.1 imply the following corollary.

Corollary 3.8.2. *Let* $P \subset \mathbb{R}^d$ *be a set of points admitting a* θ*-well-paced ordering for* $\theta \leq \frac{1}{2}$*, let* N *be a boundary net as constructed in Section 3.5, and let* τ *be a constant greater than* 2*. If* $M = \textsc{VoronoiRefine}(P, N, \tau)$*, then* $|M| = O(n \log \frac{1}{\theta})$.

3.8.1 Upper bounds on the feature size integral

The proof of the upper bound on the feature size integral will follow a simple pattern. First, we prove a bound for inputs consisting of only two points (Lemma 3.8.3). Then, we bound the change in the integral upon adding a single new point (Lemma 3.8.4). Finally, we apply this Lemma inductively to get the final bound (Lemma 3.8.5).

Lemma 3.8.3. *If* $P = \{p, q\}$ *and* $\Omega \subset \mathbf{ball}(p, c|p-q|)$ *for some constant* $c > 1$*, then*

$$\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} \leq \mathbb{V}_d(1 + \ln(2c)).$$

Proof. For all $x \in \mathbb{R}^d$, $\mathbf{f}_P(x) \geq \max\{\frac{1}{2}|p-q|, |x-p|\}$. So, we can rewrite the integral in polar coordinates (centered at p) and bound it as follows.

$$\begin{aligned}
\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} &\leq d\mathbb{V}_d \left(\int_0^{\frac{1}{2}|p-q|} \frac{r^{d-1} dr}{(\frac{1}{2}|p-q|)^d} + \int_{\frac{1}{2}|p-q|}^{c|p-q|} \frac{r^{d-1} dr}{r^d} \right) \\
&\leq d\mathbb{V}_d \left(\frac{1}{d} + \ln\left(\frac{2c|p-q|}{|p-q|} \right) \right) \\
&= \mathbb{V}_d(1 + \ln(2c)).
\end{aligned}$$

□

We now bound the change in the feature size integral induced by the addition of a single point.

Lemma 3.8.4. *Let* P *be a point set and let* $P' = P \cup \{q\}$*. If* q *is* θ*-medial with respect to* P *then*

$$\int_\Omega \left(\frac{1}{\mathbf{f}_{P'}(x)^d} - \frac{1}{\mathbf{f}_P(x)^d} \right) dx \leq \mathbb{V}_d \left(1 + d \ln \frac{3d}{\theta} \right).$$

Proof. Let U be the subset of $\mathbb{R}^d$ where $\mathbf{f}_P \neq \mathbf{f}_{P'}$. Clearly, the integral is 0 outside U, so we can restrict our attention to U. Let R be the distance from q to the nearest point of P. For all points x in the ball $B = \mathbf{ball}(q, \frac{R}{2})$, $\mathbf{f}_{P'}(x) \geq \frac{R}{2}$, so

$$\int_B \left(\frac{1}{\mathbf{f}_{P'}(x)^d} - \frac{1}{\mathbf{f}_P(x)^d} \right) dx \leq \int_B \left(\frac{2}{R} \right)^d dx = \mathbb{V}_d.$$

The definitions imply an upper bound on $\mathbf{f}_P$ and a lower bound on $\mathbf{f}_{P'}$ for any point $x \in U$:

$$\mathbf{f}_P(x) \leq |x - q| + \frac{R}{\theta}, \tag{3.6}$$

$$\mathbf{f}_{P'}(x) \geq |x - q|. \tag{3.7}$$

The upper bound follows because q is θ-medial. The lower bound follows because q must be one of the two nearest neighbors of x if $\mathbf{f}_P(x) \neq \mathbf{f}_{P'}(x)$. We apply these bounds as follows.

$$\begin{aligned}
\int_{U \setminus B} \left(\frac{1}{\mathbf{f}_{P'}(x)^d} - \frac{1}{\mathbf{f}_P(x)^d} \right) dx &\leq \int_{U \setminus B} \left(\frac{1}{|x-q|^d} - \frac{1}{(|x-q| + R/\theta)^d} \right) dx && \text{[by (3.6) and (3.7)]} \\
&\leq \int_{\mathbb{R}^d \setminus B} \left(\frac{1}{|x-q|^d} - \frac{1}{(|x-q| + R/\theta)^d} \right) dx && \text{[integrand is nonnegative]} \\
&= d\mathbb{V}_d \int_{R/2}^{\infty} \left(\frac{1}{r^d} - \frac{1}{(r + R/\theta)^d} \right) r^{d-1} dr && \text{[in polar coordinates]} \\
&< d\mathbb{V}_d \ln \frac{3d}{\theta}. && \text{[Lemma 3.10.1]}
\end{aligned}$$

The final inequality follows from a straightforward calculus exercise (the full proof may be found in Lemma 3.10.1 below). To bound the integral over all of Ω, we simply add the bounds on the integral over B and $U \setminus B$. □

Lemma 3.8.5. *Let $P = \{p_1 \ldots, p_n\}$ be an ordered set of points with prefixes $P_i = \{p_1, \ldots, p_i\}$ such that $|p_1 - p_2| = \mathbf{diameter}(P)$. Let $\Omega \subset \mathbf{ball}(p_1, c\mathbf{diameter}(P))$ for some constant $c > 1$ be the bounding region. Then,*

$$\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} < \mathbb{V}_d \left(1 + \ln(2c) + \sum_{i=3}^{n} \left(1 + d \ln \frac{3d}{\theta_i} \right) \right),$$

where $\theta_i = \frac{\mathbf{f}_{P_i}(p_i)}{\mathbf{f}_{P_{i-1}}(p_i)}$.

Proof. By definition, each point p_i is θ_i-medial with respect to P_{i-1}. By Theorem 3.4.1, it will suffice to bound the feature size integral, which may rewritten as a telescoping sum:

$$\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} = \int_\Omega \frac{dx}{\mathbf{f}_{P_2}(x)^d} + \sum_{i=3}^{n} \left(\int_\Omega \frac{dx}{\mathbf{f}_{P_i}(x)^d} - \int_\Omega \frac{dx}{\mathbf{f}_{P_{i-1}}(x)^d} \right).$$

Now, by Lemma 3.8.4 we can simplify this to get

$$\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} < \int_\Omega \frac{dx}{\mathbf{f}_{P_2}(x)^d} + \mathbb{V}_d\left(\sum_{i=3}^{n}\left(1 + d\ln\frac{3d}{\theta_i}\right)\right).$$

The base case of the telescoping sum is handled by Lemma 3.8.3 to complete the proof. □

3.8.2 Lower bounds on the feature size integral

Lemma 3.8.6. *Let P be a set of at least 2 points and let $P' = P \cup \{q\}$ for some $q \in \mathbb{R}^d$. Let $\Omega \subset \mathbb{R}^d$ be a set containing* $\mathbf{ball}(q, 2\mathbf{diameter}(P'))$. *If $\frac{\mathbf{f}_{P'}(q)}{\mathbf{f}_P(q)} = \theta$ then*

$$\int_\Omega \left(\frac{1}{\mathbf{f}_{P'}(x)^d} - \frac{1}{\mathbf{f}_P(x)^d}\right) dx \geq \mathbb{V}_d\left(\frac{d\ln\frac{1}{\theta}}{3^d} + \theta^d - 1\right)$$

Proof. Let $R = \mathbf{f}_{P'}(q)$ Let $U = \{x : \frac{R}{2} \leq |x-q| \leq \frac{R}{2\theta}\}$. For all $x \in U$,

$$\mathbf{f}_P(x) \geq \frac{R}{2\theta}, \text{ and}$$
$$\mathbf{f}_{P'}(x) \leq 3|x-q|.$$

The lower bound follows because q is θ-medial: there is at most one point of P in $\mathbf{ball}(q, \frac{R}{\theta})$ and therefore at most one point in $\mathbf{ball}(x, \frac{R}{2\theta})$. The upper bound follows because $\mathbf{f}_{P'}(x) \leq \mathbf{f}_{P'}(q) + |x-q|$ and $\mathbf{f}_{P'}(q) = R \leq 2|x-q|$ by the definition of U.

$$\begin{aligned}
\int_U \left(\frac{1}{\mathbf{f}_{P'}(x)^d} - \frac{1}{\mathbf{f}_P(x)^d}\right) dx &\geq \int_U \left(\frac{1}{(3|x-q|)^d} - \frac{(2\theta)^d}{R^d}\right) dx \\
&\geq d\mathbb{V}_d \int_{\frac{R}{2}}^{\frac{R}{2\theta}} \left(\frac{1}{(2r)^d} - \frac{(2\theta)^d}{R^d}\right) r^{d-1} dr \\
&= d\mathbb{V}_d \left(\left(\frac{1}{3^d}\int_{\frac{R}{2}}^{\frac{R}{2\theta}} \frac{dr}{r}\right) + \frac{\theta^d - 1}{d}\right) \\
&= \mathbb{V}_d\left(\frac{d\ln\frac{1}{\theta}}{3^d} + \theta^d - 1\right)
\end{aligned}$$

□

The lower bound is not informative if θ is large. This is unavoidable because the addition of a 1-medial point may have a negligibly small effect on the feature size integral. When we apply the lower bound to a set of n points, we get the following.

Lemma 3.8.7. *Let $P = \{p_1 \ldots, p_n\}$ be an ordered set of points with prefixes $P_i = \{p_1, \ldots, p_i\}$. Let $\Omega \subset \mathbb{R}^d$ be a set containing* $\mathbf{ball}(p, 3\mathbf{diameter}(P))$ *for some $p \in P$.*

$$\int_\Omega \frac{dx}{\mathbf{f}_P(x)^d} > \mathbb{V}_d \left(\frac{d}{3^d} \sum_{i=3}^{n} \ln \frac{1}{\theta_i} - n \right),$$

where $\theta_i = \frac{\mathbf{f}_{P_i}(p_i)}{\mathbf{f}_{P_{i-1}}(p_i)}$.

Proof. By definition, each point p_i is θ_i-medial with respect to P_{i-1}. Consequently, the result follows directly from Lemma 3.8.6. □

3.9 Concluding Remarks

This chapter gave a relatively complete look at bounding mesh sizes in terms of the local feature size, from generalizations of the classic Ruppert bounds to tight to per-instance bounds. We now have a good set of structural theorems to attack some of the problems that come up later in the thesis.

3.10 Technical Lemmas

Lemma 3.10.1. *Given positive constants $\theta \leq 1$, c, and R,*

$$\int_{R/c}^{\infty} \left(\frac{1}{r^d} - \frac{1}{(r + \frac{R}{\theta})^d} \right) r^{d-1} dr < \ln \frac{d(c+1)}{\theta}.$$

Proof. We bound this integral using the change of variables $u = \frac{R}{r\theta} + 1$ as follows.

$$\begin{aligned}
\int_{R/c}^{\infty} \left(\frac{1}{r^d} - \frac{1}{(r + \frac{R}{\theta})^d} \right) r^{d-1} dr &= \int_1^{1+c/\theta} \left(\left(\frac{\theta(u-1)}{R} \right)^d - \left(\frac{\theta(u-1)}{Ru} \right)^d \right) \left(\frac{R^d}{\theta^d (u-1)^{d+1}} \right) du \\
&= \int_1^{1+c/\theta} \left(1 - \frac{1}{u^d} \right) \left(\frac{1}{u-1} \right) du \\
&= \int_1^{1+c/\theta} \left(\frac{u^d - 1}{u^d (u-1)} \right) du \\
&= \sum_{i=0}^{d-1} \int_1^{1+c/\theta} u^{i-d} du \\
&= \ln \left(1 + \frac{c}{\theta} \right) + \sum_{i=0}^{d-2} \left(\frac{(1 + \frac{c}{\theta})^{i-d+1}}{i-d+1} - \frac{1}{i-d+1} \right) \\
&= \ln \left(1 + \frac{c}{\theta} \right) + \sum_{j=1}^{d-1} \left(\frac{1}{j} - \frac{1}{j(1 + \frac{c}{\theta})^j} \right) \\
&< \ln \left(1 + \frac{c}{\theta} \right) + \ln d \\
&\leq \ln \frac{d(1+c)}{\theta}.
\end{aligned}$$

□

Chapter 4

An Optimal Algorithm for Meshing Point Sets

4.1 Overview

We present NetMesh, a new algorithm that produces a conforming Delaunay mesh for point sets in any fixed dimension with guaranteed optimal mesh size and quality. Our comparison based algorithm runs in time $O(n \log n + m)$, where n is the number of input vertices and m is the number of output vertices, and with constants depending only on the dimension and the desired element quality bounds. It can terminate early in $O(n \log n)$ time returning a $O(n)$ size Voronoi diagram of a superset of P with a relaxed quality bound, which again matches the known lower bounds.

The previous best results in the comparison model depended on the log of the **spread** of the input, the ratio of the largest to smallest pairwise distance among input points. We reduce this dependence to $O(\log n)$ by employing the theory of range space ϵ-nets, a sequence of ϵ-nets determine the insertion order in an incremental Voronoi diagram. We generate a hierarchy of well-spaced meshes and use these to show that the complexity of the Voronoi diagram stays linear in the number of points throughout the construction.

4.2 Meshing point sets

In this chapter we present a new algorithm for meshing point sets in fixed dimension. This is the first algorithm we know of that is work-optimal in the comparison-based model in the sense of [Yao81]. Known work-efficient algorithms for meshing are one of two types. The first of these are based on incremental refinement of the Voronoi diagram or Delaunay triangulation. The only work-efficient of these in higher dimension performs a recursive Voronoi refinement where at all times a "quality" Voronoi mesh is maintained. Unfortunately, this leads to work of $O(n \log \Delta + m)$ where Δ is the spread of P [HMP06, HMP07]. The second type uses a quadtree to generate a mesh. Work-efficient versions use bit manipulation of the coordinates of the points to efficiently help with

the point location [MV00, BET99, HPÜ05]. These algorithms are not optimal in the comparison model and possibly more importantly, it is not known how to efficiently handle higher dimensional features (segments, facets) with these methods.

Our algorithm uses range space ϵ-nets to determine the insertion order of the input points to improve the work bound for point sets with large spread. Clarkson used a similar method for doing point location in a Voronoi diagram [Cla88]. In our approach, since we also add some Steiner points, we can guarantee that the total size of the intermediate Voronoi diagrams are only linear size. This insertion order requires us to maintain a Voronoi diagram that need not have good aspect ratio in the usual sense.

Our algorithm will generate a linear-size mesh in fixed constant dimensions. In their 1994 paper Bern, Eppstein, and Gilbert showed how to generate such a linear-size mesh with no large angles [BEG94]. In a later paper we gave a Voronoi refinement algorithm that also generates linear size meshes, [MPS08] but had a running time of only $O(n \log \Delta)$.

Because a standard good aspect ratio mesh is too large, we maintain a weaker but sufficient condition, bounded ply. Throughout the life of the algorithm we maintain a mesh that is of bounded ply which will be used to bound the point location work and the work to determine the insertion order:

Definition. *A Voronoi Diagram of a domain Ω is k-**ply** if for every point $x \in \Omega$ at most k D-balls contain x in their interior.*

Using the bounded-ply property we can afford to maintain a copy of each uninserted point in each Delaunay ball that contains it. We pick an insertion ordering so that the number of uninserted points stored in a Delaunay ball decreases geometrically, which we achieve using ϵ-nets.

Let P be the input points and M be points that have been inserted into the mesh so far including the Steiner points. We say that M is an ϵ-net for P if any ball whose interior is disjoint from M contains at most ϵn points from P (we give the formal definition of range space ϵ-nets in Section 4.10). We show, given a mesh M that is an ϵ-net, how to pick at most a constant number of points per Delaunay ball so that after their insertion the new mesh will be a $\epsilon/2$-net. Thus, a **round** consists of adding these new input points plus a constant factor more Steiner points so that we recover a bounded-ply mesh. After $O(\log n)$ rounds the process terminates with a constant ply mesh of size $O(n)$. This output can then be finished to a standard good aspect ratio mesh in output sensitive $O(m)$ time if desired.

4.3 Beating the Spread

The spread of a point set is the ratio of the largest to smallest interpoint distances, and is denoted as Δ. It is a (geo)metric rather than a combinatorial property; given a set of points P, its cardinality may be n but its spread is not in general bounded by any function of n. It is not uncommon to see a dependence on the spread in the analysis of algorithms in computational geometry and finite

metric spaces. Though rarely a problem in practice, it does thwart the most basic principle in the analysis of algorithms, to bound the complexity in terms of the input size.[1]

Consider two classic data structures, the quadtree and the kd-tree. The quadtree partitions space geometrically, breaking squares into 4 pieces of equal **size**. The kd-tree partitions the input points combinatorially into sets of equal **cardinality**. These data structures demonstrate the difference between geometric and combinatorial divide and conquer. The quadtree has depth $\log \Delta$ whereas the kd-tree has depth $\log n$. Unfortunately, many computational problems from nearest neighbor search to network design problems depend on (geo)metric information that is lost when doing a combinatorial divide and conquer. Thus, for many problems, the best known algorithms depend on the spread in either time or space complexity or both.

One approach to dealing with the spread is to restrict the computational model. If coordinates are restricted to be $\log n$-bit integers then the spread is $O(n)$. If we use floating point numbers, the spread is $O(2^n)$. These assumptions about the bit representation of the input also allow for fast computation of logarithms as well as the floor and ceiling functions. These computations are usually omitted from the basic operations of the real RAM model often used in computational geometry to extend the comparison sorting model from the real line to d-dimensional Euclidean space. In their work on metric nets, Har-Peled and Mendel correctly argue that if one can do arithmetic in constant time, it is natural to expect also to perform other operations of size $O(\log \log \Delta)$ in constant time [HPM06]. This is certainly the case for many practical implementations of geometric algorithms. However, it is interesting, both in theory and in practice to explore ways of eliminating the dependence on the spread without resorting to specialized bit operations–in theory because it probes the limits of an important computational model and in practice because it allows one to work with a minimal set of primitives with minimal assumptions about the low-level data representation.[2]

In mesh generation, a dependence on the spread creeps in from two different sources, in the output size and the in the cost of point location. The previous state of the art in comparison based point meshing requires $O(n \log \Delta + m)$ work, where the first term is the cost of point location and the second is the output sensitive term. Even for point set inputs, the lower bounds on quality meshes imply that m may also depend on the spread. Thus, to avoid depending on the spread, we must both optimize point location and also relax the quality condition. This is why our algorithm has two phases, one that produces a linear size Voronoi diagram of a superset in $O(n \log n)$ time and one that refines that mesh to quality in $O(m)$ time.

4.4 Sparse Refinement and Point Location

There are two immediate challenges to time-optimal meshing. The first challenge is the complexity of general Voronoi diagrams, which can be as large as $\Omega(n^{\lceil d \rceil})$. This challenge was met by the Sparse

[1] One can get around this by making assumptions about the bit representations of the inputs. We will address this as well.

[2] Recall that in the popular CGAL library, all primitives are implemented for several different kernels, all use a small, unified interface.

Voronoi Refinement (SVR) algorithm of Hudson et al. [HMP06]. Their algorithm guarantees that the complexity of the Voronoi diagram remains linear in the number of vertices at every stage of the algorithm. The second challenge is to efficiently handle point location; to add a point to a Voronoi diagram, one must first locate that point in the current Voronoi diagram. This is where SVR falls short of optimality.

In this section, we give an overview of how SVR avoids the worst case complexity of Voronoi diagrams. Then we explain the basic point location paradigm used by both SVR and the NetMesh algorithm to be presented in Section 4.7. This will give a foundation for understanding the new ideas introduced in NetMesh to achieve optimality.

Sparse Voronoi Refinement. Recall that in the VoronoiRefine algorithm defined in Section 3.2, all of the input points are added prior to the insertion of any Steiner points. The key insight of SVR is to interleaves the insertion of input points and Steiner points. In doing so, the algorithm requires two extra pieces. First, input points are only added if they are "close" to the current Voronoi diagram. Second, the Steiner points may not be added "too close" to uninserted input points. The former notion of closeness is ε-mediality (see page 10). The latter notion of closeness causes the algorithm to **yield** by adding an input point p rather than a Steiner point v if the distance from p to v is less than γ times the radius of the empty ball around v.

By only inserting ε-medial points and yielding when appropriate, SVR maintains a good aspect ratio Voronoi diagram at every stage of the algorithm. Consequently, the total work is output sensitive. This approach has also been generalized to more complex inputs than just point sets, considering also piecewise linear complexes [HMP06].

Point Location. The bottleneck for the running time of Voronoi refinement is point location. Recall, that in the standard incremental Voronoi (or Delaunay) algorithm, the first step to inserting a new point is to find that point in the current diagram. A natural and highly effective technique for doing this point location is to eagerly store the uninserted points in the D-balls of each Voronoi diagram as the algorithm progresses. Points are moved whenever an insertion changes a D-ball locally.

In SVR, this approach corresponds to a geometric divide and conquer, similar in spirit to quadtree methods, because after a constant number of moves, the size (radius) of the balls containing any point goes down by a constant factor. Thus, in SVR a single input point may be moved $\Theta(\log \Delta)$ times. In this chapter, we show how to modify the algorithm so that only $O(\log n)$ moves are necessary. One way to view these results is as a way to achieve similar properties to compressed quadtrees without leaving the comparison model or privileging any fixed set of coordinate axes.

4.5 Hierarchical Meshes and Hierarchical Quality

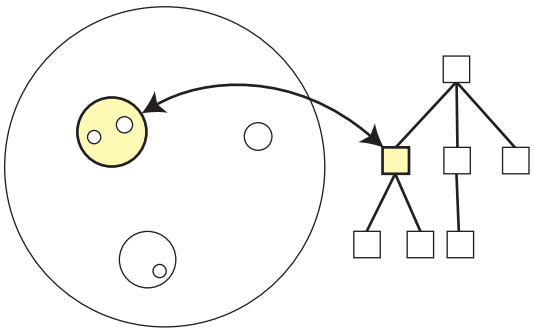

Figure 4.1: A domain hierarchy as a collection of sets (left) and its tree structure (right).

Domains. A **domain** $\Omega \subset \mathbb{R}^d$ is defined by a center c_Ω, a radius r_Ω, and a collection of disjoint open balls $B_1, \ldots, B_k \subset B_\Omega = \textbf{ball}(c_\Omega, r_\Omega)$ such that

$$\Omega = B_\Omega \setminus \left(\bigcup_{i=1}^{k} B_i \right).$$

The ball B_Ω is called the **bounding ball** of Ω and $S_\Omega = \partial B_\Omega$ is the **bounding sphere** of Ω.

We get a hierarchy of domains if the balls removed from B_Ω are the bounding balls of other domains. Formally, a **domain hierarchy** is set H of disjoint domains with a unique parent $\mathrm{p}(\Omega)$ of each $\Omega \in H$ except for a unique root domain Ω_{root} such that:

1. for any pair $\Omega, \Omega' \in H$, $\mathrm{p}(\Omega') = \Omega$ if and only if $S_{\Omega'} \subset \Omega$, and
2. $\bigcup_{\Omega \in H} \Omega = B_{\Omega_{\text{root}}}$.

The parent relation induces a tree structure on the domains of H with root at Ω_{root}. The set of all domains $\Omega' \in H$ such that $\mathrm{p}(\Omega') = \Omega$ is denoted $\mathrm{children}(\Omega)$.

Cages. Given a domain Ω, we want to add vertices near S_Ω to limit the interaction between the inside and the outside of Ω. We will have two parameters, δ determining the density of these points, and γ determining how nearly cospherical they are. We call such a set C_Ω of vertices a **cage** and we require the following three properties, where $\overline{r} = (1 - \delta - \gamma) r_\Omega$ and $\overline{S} = (1 - \delta - \gamma) S_\Omega$.

1. **[Nearness Property]** For all $v \in C_\Omega$, $\mathbf{d}(v, \overline{S}) \leq \gamma \overline{r}$.
2. **[Covering Property]** For all $x \in \overline{S}$, $\mathbf{d}(x, C_\Omega) \leq (\delta + \gamma)\overline{r}$.
3. **[Packing Property]** For all distinct $u, v \in C_\Omega$, $|u - v| \geq (\delta - 2\gamma)\overline{r}$.

These three properties are illustrated in Figure 4.2

To construct such a set of points, we start with a **cage template** T of points on the unit sphere S. The points of T are a metric space δ-net on S (not to be confused with the range space nets used elsewhere in this chapter). That is, for all $x \in S$, $\mathbf{d}(x, T) \leq \delta$ and for each distinct pair

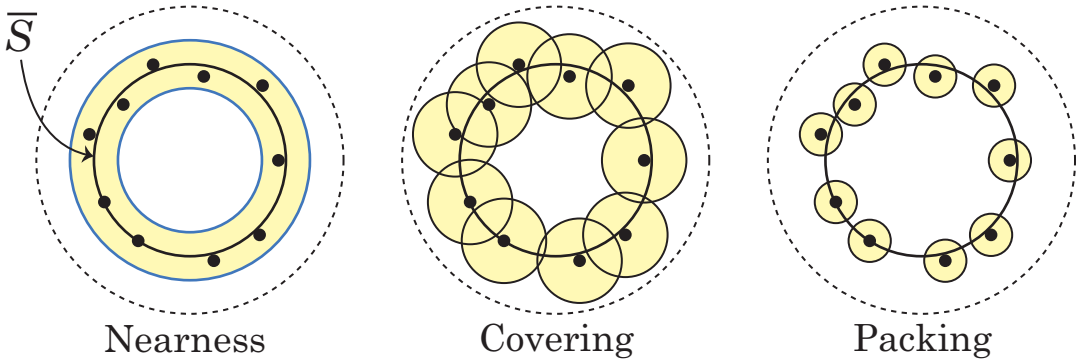

Figure 4.2: The three properties of the cage vertices.

$u, v \in T$, $|u - v| \geq \delta$. Such sets are known to exist and can be constructed using a simple greedy algorithm [Gon85, Mat02].

For a domain Ω we construct its cage by adding for each $x \in c_\Omega + \overline{r}T$, a new point x' such that $|x - x'| \leq \gamma\overline{r}$. It is easy to check that this set of points will satisfy the three properties of a cage.

Definition. *A cage C_Ω centered at c with radius r is ε-**encroached** or simply **encroached** by a point $p \notin C_\Omega$ if either*

1. *p is an input point in* $\mathbf{annulus}(c, \varepsilon r, r)$*, (**inner-encroachment**), or*
2. *p is an input or Steiner point in* $\mathbf{annulus}(c, r, \frac{2r}{\varepsilon})$*, (**outer-encroachment**).*

Roughly speaking, non-encroached cages have room on the inside (w.r.t. input points) and room on the outside (w.r.t. all mesh vertices).

Hierarchical Meshes.

Definition. *A **hierarchical mesh** is a mesh M along with a domain hierarchy H_M such that:*

1. *M has a vertex at the center of every domain, i.e. $c_\Omega \in M$ for all $\Omega \in H_M$*
2. *No domain is ε-encroached.*

Given a hierarchical mesh M and $\Omega \in H_M$, we define M_Ω to be the points of M contained in Ω plus the centers of the children of Ω in H_M. Formally,

$$M_\Omega = (M \cap \Omega) \cup \{c_{\Omega'} : \Omega' \in \text{children}(\Omega)\}.$$

We call this the set M restricted to the domain Ω, and it is well defined for any domain Ω and any set M that contains the centers of the children of Ω. In particular, for a subset $P \subset M$, $P_\Omega = P \cap M_\Omega$.

In a hierarchical mesh, we can also define the Voronoi cell of a cage C_Ω as

$$\text{Vor}_M(C_\Omega) = \bigcup_{u \in M \cap B_\Omega} \text{Vor}_M(u).$$

Definition. *We say that a hierarchical mesh M is τ-**quality** if the following conditions are met:*

1. *For every non-cage vertex $v \in M$, $\mathrm{Vor}_M(v)$ has aspect ratio at most τ.*
2. *For every $\Omega \in H_M$, $\mathrm{Vor}_M(C_\Omega)$ has aspect ratio at most τ.*
3. *No domain in H_M is ε-encroached.*

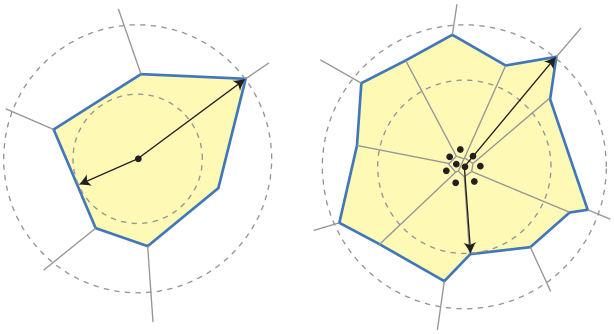

Figure 4.3: Quality cells of a vertex (left) and a cage (right).

The four constants γ, δ, ε, and τ are called the **meshing parameters**. Throughout, they are assumed to be fixed constants independent of the dimension.

Definition. *For a set M and a domain Ω, the **feature size** is a function $\mathbf{f}_M^\Omega : \mathbb{R}^d \to \mathbb{R}$ that maps a point x to the distance to its second nearest neighbor among the points of M_Ω.*

We are mainly interested in the feature size of the input and of the mesh, $\mathbf{f}_P^\Omega$ and $\mathbf{f}_M^\Omega$ respectively, over the domains of M_H.

4.6 Additively-Weighted Voronoi Diagrams

There is a natural generalization of Voronoi diagrams in which the points are permitted to have weights that affect the distance additively. For a point v, let w_v be the weight of v. The distance between two weighted points is defined as

$$\mathbf{d}_w(u, v) = |u - v| - w_u - w_v.$$

Given a finite set of weighted points M, this distance is extended to all of $\mathbb{R}^d$ by assuming that $w_x = 0$ for all $x \in \mathbb{R}^d \setminus M$. An equivalent formulation simply measures the distance between the spheres of radius w_u and w_v centered at u and v respectively. For a point set M, we define the **additively-weighted Voronoi cell** of a point $v \in M$ to be

$$\mathrm{Vor}(v) = \{x \in \mathbb{R}^d : \min_{u \in M} \mathbf{d}_w(u, x) = \mathbf{d}_w(v, x)\}.$$

The **additively-weighted Voronoi diagram** is the cell complex decomposing $\mathbb{R}^d$ obtained by taking all of the additively-weighted Voronoi cells. This definition generalizes the standard Voronoi diagram, which may be viewed as an additively-weighted Voronoi diagram for points of 0 weight.

The additively-weighted Voronoi diagram is different from the more common notion of weighted Voronoi diagrams obtained by replacing the distance function with the power distance. The cells of the additively-weighted diagram are not necessarily polyhedra nor are they necessarily convex. Still, it is possible to extend basic properties of Voronoi diagrams to the case of additive weights. The **in-radius** of a Voronoi cell $\mathrm{Vor}(v)$ is defined as

$$\textbf{in-radius}(\mathrm{Vor}(v)) := \max\{r : \textbf{ball}(v, r) \subset \mathrm{Vor}(v)\},$$

and similarly, the **out-radius** is defined as

$$\textbf{out-radius}(\mathrm{Vor}(v)) := \min\{r : \mathrm{Vor}(v) \subset \textbf{ball}(v, r)\}.$$

Note that $\textbf{ball}(v, r)$ in the above definitions is a Euclidean ball, not a ball with respect to the weighted distance $\mathbf{d}_w$. We say that u and v are neighbors if $\mathrm{Vor}(u) \cap \mathrm{Vor}(v) \neq \emptyset$. The in-radius of $\mathrm{Vor}(v)$ may also be defined as $\frac{1}{2}\mathbf{d}_w(u, v) + w_v$, where u is the nearest among the neighbors of v in additive distance. The **aspect ratio** of $\mathrm{Vor}(v)$ is the ratio of the out-radius to the in-radius.

Approximation by Cages. The additively-weighted Voronoi diagram can be approximated by a regular Voronoi diagram by replacing the weighted points with a small cage of new vertices at distance w_v from each weighted point v. The approximate cells are the union of the Voronoi cells of the cage vertices. These approximate cells can also be used to get a good approximation of the in-radius and out-radius of the weighted Voronoi cell. Let v be a vertex with cage vertices C. The neighbors of C are those vertices v that share a Voronoi face with a vertex in C but are not in $C \cup \{v\}$. The in-radius of the approximate Voronoi cell is $\frac{1}{2}\mathbf{d}_w(u, v) + w_v$, where u is the nearest among the neighbors of C in additive distance. Since the neighbors necessarily have weight 0, this reduces to $\frac{|v-u|+w_v}{2}$.

Recall that for unweighted points M, the feature size function $\mathbf{f}_M : \mathbb{R}^d \to \mathbb{R}$ is the distance to the second nearest point of M. So, in the absence of weights, the in-radius of $\mathrm{Vor}(v)$ is $\frac{1}{2}\mathbf{f}_M(v)$ and if the aspect ratio is τ and the out-radius is R then $\mathbf{f}_M(v) = \frac{2R}{\tau}$. If the points have weights then the definition of $\mathbf{f}_M$ is the same as if the points have no weights. The following lemma shows how the feature size relates to the out-radius and aspect ratio of the weighted Voronoi cells.

Lemma 4.6.1. *Let v be a point in a weighted set M. Let r be the in-radius of* $\mathrm{Vor}(v)$*. If $w_a \leq \varepsilon(|a - b| - w_b)$ for all $a, b \in M$, then*

$$\frac{2r(1-\varepsilon)}{1+\varepsilon} \leq \mathbf{f}_M(v) \leq \frac{2r}{1-\varepsilon}$$

Proof. Let x and y be the nearest points to v in Euclidean and weighted distance respectively (it could be that $x = y$). So, $2r = |v - y| + w_v - w_y$ and $\mathbf{f}_M(v) = |x - v|$. By assumption, w_v and w_y

are both less than $\varepsilon|v-y|$. So, it follows that

$$(1-\varepsilon)|v-y| \leq |v-y| + w_v - w_y \leq (1+\varepsilon)|v-y|. \tag{4.1}$$

This assumption and the definitions of u and y also imply that

$$(1-\varepsilon)|v-y| \leq |x-v| \leq |v-y|. \tag{4.2}$$

So, the result follows from (4.1) and (4.2). □

When we choose sufficiently dense cages, nearly the same bounds apply for the approximate weighted Voronoi cells:

Lemma 4.6.2. *Let v be a vertex or a cage in a hierarchical mesh M. Let r be the in-radius of* $\mathrm{Vor}(v)$ *and let c be the center of v. If no cages are ε-encroached for ε sufficiently small, then*

$$r \leq \mathbf{f}_M^{\Omega}(c) \leq 3r,$$

where Ω is the domain containing the boundary of $\mathrm{Vor}(v)$.

4.7 The Algorithm

4.7.1 Overview of the Algorithm

Like SVR, the core of the NetMesh algorithm is an incremental construction of a Voronoi diagram with the refinement steps to maintain mesh quality. There are five main concerns. The algorithm must **(1) order the input points**. These points are added one at a time in an **(2) incremental construction**. After each insertion, Steiner points are added in a **(3) refinement** phase that recovers the quality invariant. All the while, uninserted points are organized in a **(4) point location** data structure. Once all of the inputs have been added, an optional **(5) finishing** procedure turns the linear-size hierarchical mesh into a standard well-spaced mesh. Each of these concerns will be addressed in more detail below, but first we will describe the main ideas used and how they fit together.

Point Location. The point location data structure associates each point with each D-ball that contains it. So, it is easy to report the set of D-balls containing an input point and similarly, to report the set of points in a D-ball. These associations are updated locally every time a new point changes the underlying Delaunay triangulation. We will prove that no point is ever in more than a constant number of D-balls and thus the size of this structure will not exceed $O(n)$.

Incremental updates. In SVR, every insertion is medial. This is critical to maintain quality in the mesh throughout the algorithm. In the NetMesh algorithm, we change the domain hierarchy before inserting each point to guarantee that it is medial in whatever domain contains it. We show that this is sufficient to get the same guarantees as in SVR. Thus, we can insert the points in any order.

Ordering the input with ϵ-nets. The theory of range space ϵ-nets is used to choose the input insertion order. One round of the algorithm consists of the union of a collection of ϵ-nets for the input points for each D-ball, where the ranges are open balls. It is known that such sets exist, are small, and can be found quickly and deterministically [Cha00]. The points in any round may be inserted in any order, after which, the next round is computed. In each round, the maximum number of points stored in any D-ball goes down by a constant factor, so the total number of rounds is $O(\log n)$.

Refinement. The refinement, or cleaning phase of the algorithm is a standard Voronoi refinement in that it adds Steiner points at the farthest corner of any cell with bad aspect ratio. As in SVR, if the Steiner point is sufficiently close to an uninserted input point p, then p is added instead. One slight change is that we maintain the aspect ratio of the Voronoi cells of cages, but do not require the cage vertices themselves to have good aspect ratio Voronoi cells.

Finishing the mesh. The algorithm produces a quality hierarchical mesh of linear size. If one wants to extend this mesh to a standard well-spaced mesh, it is a straightforward procedure to do this in $O(m)$ time, where m is the number of vertices in an optimal-size, well-spaced superset of P. This finishing process can run quickly because it need not do any point location (all of the input points have already been inserted).

4.7.2 Point Location Operations

Each uninserted input point p stores a list of D-balls that contain it as well as a list of cages that it encroaches. The set of all D-balls in a mesh M is DBALLS(M). Similarly, each D-balls B has a list of uninserted vertices that it contains, UNINSERTED(B). With each change in the Voronoi diagram, these lists are updated. We say that the points are "stored *in* the balls" to simplify the description of this list upkeep. A point will generally be contained in several D-balls. The uninserted points are moved out of D-balls that have been destroyed and into newly created D-balls. This shuffling of points between D-balls is the work of point location. A point is touched in this process if it is moved into a new ball or even if it is considered for moving into a new ball. We count the point location work from the perspective of the uninserted input points.

There are four main point location operations needed.

1. Find the D-balls containing a point to insert it into the Voronoi diagram.
2. Find the nearest and second nearest neighbor of a point in its domain in order to compute its mediality, which is the ratio of these distances.
3. Find any cages encroached by a given point.
4. Find a nearby input point to yield to, when inserting a Steiner point.

The first operation is trivial.

For cage vertices v in a domain Ω, let $center(v)$ be the vertex at the center of Ω. Let $B(x)$ be the set of D-balls containing x. Let $V(B)$ be the $d+1$ vertices of the Delaunay simplex corresponding to the D-ball B. Let $U(x) = \bigcup_{B \in B(x)} V(B)$. If Ω is the domain containing x, then the nearest and

second nearest neighbors of x in M_Ω are in $U(x)$ or $\{center(v) : v \in U(x)\}$, so it is easy to identify them. Call these vertices n_x and s_x respectively. Thus, MEDIALITY$(x) = \frac{|x-n_x|}{|x-s_x|}$ can be computed in time $O(|B(x)|)$.

The set of cages outer-encroached by a point x is denoted OUTENCROACH(x). Similarly, the set of cages inner-encroached by x is INENCROACH(x). Computing these sets for input points is easy because this information is stored with the points. At the time a cage is created, any encroaching input points must be relocated, so the encroachment is discovered at that time. To check encroachment of Steiner points, it suffices to observe that if a Steiner point x encroaches a cage C, then some vertex of C must appear in $U(x)$. So, there are only $O(|U(x)|)$ cages to check and each check takes constant time.

To find a point to yield to, we simply need to check for input points in a small empty ball around the proposed input point. This is trivial for Steiner points added during refinement because the Steiner point is the center of a D-ball B and thus we only need to check UNINSERTED(B). For cage vertices v, the search requires us also to check the points in UNINSERTED(B) for each $B \in B(v)$. In both cases, the points checked in this process also need to be checked for relocation when the new vertex is inserted. Thus, the cost for this search is dominated by the cost of relocating points, which we analyze in detail later.

4.7.3 Incremental Updates to Hierarchical Meshes

The basic operation in incremental Voronoi diagrams is INSERT(v), which adds the vertex v to the Voronoi diagram and updates the point location data structures. To keep this operation constant time (not counting the cost of point location), we must guarantee that $|B(v)|$ is a constant because every D-ball in $B(v)$ is destroyed by the insertion. This is done by making sure that every new insertion is medial. Before the new point is inserted, we update the domain hierarchy. If the point was not medial, then it must be significantly closer to its nearest neighbor than it second nearest neighbor, and thus we add or expand cage around the nearest neighbor. We must also update the domain hierarchy of the new point encroaches on an existing cage.

INSERT$(x \in \mathbb{R}^d)$

for each C in OUTENCROACH(x): RELEASECAGE(C)

Add x to Vor_M and update the point location structure.

YIELDINGINSERT$(x \in \mathbb{R}^d)$

let v be the nearest neighbor of x in the current mesh.

if there is an input point p in **ball**$(x, \gamma|x - v|)$

then INSERT(p) **else** INSERT(x)

There are three basic cage operations:

NewCage($p \in P, r \in \mathbb{R}$)
Initialize a new cage.
for each $x \in T$, YieldingInsert($rx + p$).

ReleaseCage(C : cage)
for each cage vertex v in C push v to the RefineList.
Delete the cage C

GrowCage(C : cage)
Let x be the center of C and let r be its radius.
if in-radius(Vor(C)) $\geq \frac{r}{\varepsilon^2}$ **then** NewCage($x, \frac{r}{\varepsilon}$).
ReleaseCage(C).

Equipped with the cage operations, we define the following routine. Its purpose is to rearrange the domain hierarchy by creating or growing new cages so that a new vertex v can be added to a domain in which it is medial.

InsertInput($p \in P$)
let u be the nearest neighbor of v in M_Ω
if Mediality(p) $\leq \varepsilon$ **then** NewCage($u, |u - p|/\varepsilon$)
for each C in InEncroach(p): GrowCage(C)
Insert(p)

4.7.4 Refinement

The algorithm maintains a list of cells with bad aspect ratio called RefineList. The cleaning procedure goes through this list and refines these cells until none are left. The RefineList is updated every time a Voronoi cell changes. The structure of the Voronoi diagram makes it easy to check the aspect ratio of a cell and Theorem 4.8.1 implies that this can be done in constant time. If a cell's aspect ratio was good but goes bad, it is added to the list. If its aspect ratio was bad but becomes good, it is removed from the list.

Clean(M: Mesh)
while RefineList is not empty
let $v \in$ RefineList
let x be the farthest corner of Vor(v) from v
YieldingInsert(x)

4.7.5 Input Ordering with ϵ-Nets

We employ the theory of range space ϵ-nets to order the inputs for insertion. The following is a special case of Theorem 4.6 from [Cha00] when the range space is defined by open balls.

Theorem 4.7.1. *Let $P \subset \mathbb{R}^d$ be a set of n points and let $\epsilon \in (0,1)$. There exists an algorithm* NET(ϵ, P) *that runs in $O(\frac{1}{\epsilon^2}(\log\frac{1}{\epsilon})^{d+1}n)$ time and returns a subset $N \subseteq P$ such that $|N| = O(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ and any open ball that contains ϵn points of P also contains a point of N. In particular, for constant ϵ the running time is linear and the net is of constant size.*

Using the NET algorithm as a black box, we select the next round of points to insert as follows.

SELECTROUND(M: mesh)
  $N \leftarrow \emptyset$
  for each $B \in$ DBALLS(M)
    $N \leftarrow N \cup$ NET$(\frac{1}{2d}$, UNINSERTED$(B))$
  return N

If the maximum number of uninserted points in a D-ball of some mesh is k, then after adding the points chosen by SELECTROUND, this maximum is at most $\frac{k}{2}$. This follows from the fact that every new D-ball is covered by at most d of the old D-balls (see Theorem 4.10.1). So, the total number of rounds is at most $\lceil \log n \rceil$. We can now give the main loop of the algorithm.

NETMESH(P: points)
  Initialize an empty mesh M
  UNINSERTED $\leftarrow P$
  let c, r be such that $P \subset$ **ball**(c, r)
  OUTERCAGE = NEWCAGE$(c, \frac{r}{\varepsilon})$
  while UNINSERTED is not empty
    $V =$ SELECTROUND(M)
    for each $v \in V$
      INSERTINPUT(v)
      CLEAN(M)
  return M

4.7.6 Finishing the Mesh

The output of NETMESH is a quality hierarchical mesh. If the desired output is a well-spaced mesh according to the traditional definition, i.e. quality with a single domain, then some finishing procedure is required. Fortunately, it is trivial given the cage operations defined above:

FinishMesh(M: Mesh)
  while there exists a cage C other than OuterCage
    GrowCage(C)
    Clean(M)

Since the cages are not encroached, they have some space around them. The FinishMesh procedure simply grows the cages until this space is filled. No new cages are formed and no point location work on input points is required.

Note that finishing the hierarchical mesh in this way may result in a mesh with more than a linear number of points because well-spaced meshes are subject to potentially superlinear (or even superpolynomial!) lowerbounds. This is why we consider the finishing operation to be optional.

4.8 Overview of the Analysis

An intermediate mesh, M_i, is the mesh after i vertices or cages have been inserted during the incremental construction. To analyze the NetMesh algorithm, we will prove that two invariants are maintained for each intermediate mesh: **the feature size invariant** and **the quality invariant**.

Definition. *A hierarchical mesh M of an input set P satisfies the **feature size invariant** if for all domains $\Omega \in H_M$ and all vertices $v \in M_\Omega$*

$$\mathbf{f}_P^\Omega(v) \leq K f_M^\Omega(v),$$

where K is a constant that depend only on the mesh parameters.

Definition. *A hierarchical mesh M satisfies the **quality invariant** if each intermediate mesh M_i is τ'-quality for some constant τ' depending only on the meshing parameters.*

The quality invariant is useful because of several properties of quality meshes.

Theorem 4.8.1. *If M is a τ-quality mesh, then*

1. *no point of $\mathbb{R}^d$ is contained in more than $O(1)$ D-balls,*
2. *no D-ball intersects more than $O(1)$ other D-balls, and*
3. *no vertex of M has more than $O(1)$ Delaunay neighbors.*

These structural results about quality meshes are known for the case of a single domain [MTTW99, HMP06]. To extend them to the case of a quality hierarchical meshes follows the same methods as in previous work. The three conclusions are proven in Theorems 4.11.5, 4.11.9, and 4.11.7 respectively.

Over a single domain, the results of Chapter 3 show that the feature size invariant suffices to prove that the number of vertices is bounded (up to constants) by the feature size integral:

$$\int_{\Omega} \frac{dx}{\mathbf{f}_P^{\Omega}(x)^d}.$$

In previous work [MPS08], we showed that the feature size integral is $O(n)$ when the input points satisfy a certain spacing condition. We prove that in each domain Ω of the hierarchy, the points of P_Ω satisfy this spacing condition (Lemma 4.9.3), allowing us to prove that the total output size is $O(n)$ (Theorem 4.9.4).

Theorem 4.8.1 and the quality invariant imply that the cost to update the Voronoi diagram for a single insertion is constant. That is, the number of combinatorial changes to the Voronoi diagram is constant for each insertion. Thus, since the total number of points added is $O(n)$, the total work is $O(n)$, not counting the cost of point location.

To bound the cost of point location, we first show that at most a constant number of vertices are added to any D-ball in the course of a round (Lemma 4.13.2). This is then used to show that the total amount of point location work is $O(n)$ per round. Since there are only $O(\log n)$ rounds, the total work is $O(n \log n)$ as desired.

Finally, in Section 4.14, we show that the FinishMesh procedure runs in $O(m)$ time. This allows us to conclude the following theorem about the overall running time.

Theorem 4.8.2. *Given n points $P \subset \mathbb{R}^d$, the* NetMesh *algorithm produces a hierarchical quality mesh of size $O(n)$ in $O(n \log n)$ time. If this is followed by the* FinishMesh *procedure, the output is a well-spaced mesh of size $O(m)$ in $O(n \log n + m)$ time.*

4.9 Size Bounds

In this section we will show that the output of NetMesh has linear size. The analysis will follow a straightforward strategy. We will argue that the algorithm never inserts a vertex too close to an existing vertex. This is known as the **insertion radius invariant**, and it allows us to prove that the **feature size invariant** holds for all intermediate meshes. We use this to prove that for all domains Ω, M_Ω has size linear in $|P_\Omega|$ from which the overall bound follows. This strategy is not new; it parallels closely the approach of Ruppert [Rup95] for Delaunay refinement and its sparse version introduced by Hudson, Miller, and Phillips [HMP06]. We have adapted it to the case of hierarchical meshes.

We say that a hierarchical mesh M is constructed **incrementally** if the vertices are added one at a time and the domains are adjusted before every insertion so that no domain is encroached. In particular, the algorithm given is such an incremental construction. The intermediate mesh after i points and cages have been added is denoted by M_i, its domain hierarchy H_{M_i} is denoted H_i, and $P_i = P \cap M_i$ is the set of inputs inserted thus far. Define the **insertion radius** of the ith vertex added v as $\lambda_v = \mathbf{f}_{M_i}^{\Omega}(v)$, where $\Omega \in H_i$ is the domain into which v was inserted.

Definition. *A hierarchical mesh M of an input set P constructed incrementally satisfies the* ***insertion radius invariant*** *if for all domains $\Omega \in H_i$ for all i and all vertices $v \in M_{i\Omega}$*

$$\mathbf{f}^{\Omega}_{P_i}(v) \leq \begin{cases} K'_C \lambda_v & \text{if } v \text{ is inserted as a cage vertex,} \\ K'_S \lambda_v & \text{if } v \text{ is inserted as a circumcenter, and} \\ K'_I \lambda_v & \text{if } v \text{ is inserted as an input vertex} \end{cases}$$

where K'_C,K'_S, and K'_I are constants that depend only on the mesh parameters.

The following lemma states that as long as the insertion radius of every vertex is not too small then the distance to its nearest neighbor is also not too small. Its proof is straightforward and reserved for the appendix.

Lemma 4.9.1. *If M is a hierarchical mesh constructed incrementally that satisfies the insertion radius invariant, then M also satisfies the feature size invariant.*

Lemma 4.9.1 implies that in order to prove that the spacing of the points in the final mesh is good, it will suffice to show that the algorithm maintains the insertion radius invariant throughout. This is proven in the following lemma.

Lemma 4.9.2. *The hierarchical mesh M constructed by the* NetMesh *algorithm satisfies the insertion radius invariant.*

Proof. We proceed by induction on the total number of vertices added. Let v be the ith vertex added and let Ω be the domain it is inserted into.

Case 1: v is a cage vertex. Since P_Ω contains at least the center of Ω, the feature size is bounded as $\mathbf{f}^{\Omega}_{P}(v) \leq r_\Omega$. By construction, adjacent cage vertices are at least αr_Ω apart, where $\alpha = (\delta - 2\gamma)(1 - \delta - \gamma)$. So, $\lambda_v \geq \alpha r_\Omega$. Combining these two facts and choosing $K'_C \geq \frac{1+\varepsilon}{\alpha}$ yields $\mathbf{f}^{\Omega}_{P_i}(v) \leq K'_C \lambda_v$ as desired.

Case 2: v is a clean move. Steiner points are added when some vertex (or cage) $u \in M_{i-1\Omega}$ has aspect ratio greater than τ. Let V_u denote this poor aspect ratio cell. Let w be the nearest neighbor of u in M_Ω, so $\mathbf{f}^{\Omega}_{M}(u) = |u - w|$. In case we yielded in order to insert v, let v' be the true circumcenter that we tried to insert. The yielding condition guarantees that

$$|v - v'| \leq \gamma |u - v'|. \tag{4.3}$$

Since u or w or both can be the center of a child domain of Ω, we need to also consider vertices u', w' of M that define the insertion radius of v and the in-radius of V_u respectively. Since w does not encroach a domain at u and $|u - w| \leq |u - v'|$, it follows that

$$|u - u'| \leq \varepsilon |u - v'|. \tag{4.4}$$

The D-ball centered at v' has radius $|u' - v'|$ and is empty of vertices, so $\lambda_v \geq |u' - v'| - |v - v'|$.

Using the triangle inequality, (4.3), and (4.4), we can bound the insertion radius as follows.

$$|u - v'| \leq \beta\lambda_v. \tag{4.5}$$

where, $\beta = \frac{1}{1-\varepsilon-\gamma}$. Since w is closer to u than v, $\mathbf{f}^{\Omega}_{M_{i-1}}(u) = \mathbf{f}^{\Omega}_{M_i}(u)$ and $\mathbf{f}^{\Omega}_{P_{i-1}}(u) = \mathbf{f}^{\Omega}_{P_i}(u)$. So, we can use induction and Lemma 4.9.1 to get that

$$\mathbf{f}^{\Omega}_{P_i}(u) \leq K'_I|u - w|. \tag{4.6}$$

We use K'_I because it is the largest of the K' constants.

We may now derive a bound on $\mathbf{f}^{\Omega}_{P_i}(v)$ as follows.

$$\begin{aligned}
\mathbf{f}^{\Omega}_{P_i}(v) &\leq \mathbf{f}^{\Omega}_{P_i}(u) + |u - v| && \left[\mathbf{f}^{\Omega}_{P_i} \text{ is 1-Lipschitz}\right] \\
&\leq \mathbf{f}^{\Omega}_{P_i}(u) + |u - v'| + |v' - v| && [\text{triangle inequality}] \\
&\leq \mathbf{f}^{\Omega}_{P_i}(u) + (1+\gamma)|u - v'| && [\text{by (4.3)}] \\
&\leq K'_I|u - w| + (1+\gamma)|u - v'| && [\text{by (4.6)}] \\
&\leq \left(\frac{3K'_I}{\tau} + 1 + \gamma\right)|u - v'| && [V_u \text{ aspect ratio} > \tau] \\
&\leq \left(\frac{3K'_I}{\tau} + 1 + \gamma\right)\beta\lambda_v && [\text{by (4.5)}]
\end{aligned}$$

So, setting $K'_S \geq \left(\frac{3K'_I}{\tau} + 1 + \gamma\right)\beta$ yields the desired bound.

Case 3: v is an input. Choose u such that $\lambda_v = |u - v|$ and let j and Ω_j be the time that u was inserted and the domain it was inserted into respectively. If $u \in C_\Omega$, then v encroaches on Ω, which is impossible. If u is an input vertex then $\lambda_v = f^{\Omega}_P(v)$ so we are done. So, we may assume that u is a Steiner point, inserted either as either a circumcenter or as a cage vertex that was later released.

We define $K'_u = K'_S$ in the former case and $K'_u = K'_C$ in the latter. By choosing $K'_I \geq \frac{K'_u}{\gamma} + 1$, we can now derive the following bound.

$$\begin{aligned}
\mathbf{f}^{\Omega}_{P_i}(v) &\leq \mathbf{f}^{\Omega}_{P_i}(u) + |u - v| && \left[\mathbf{f}^{\Omega}_{P_i} \text{ is 1-Lipschitz}\right] \\
&\leq \mathbf{f}^{\Omega_j}_{P_j}(u) + |u - v| && [\text{by Lemma 4.17.1}] \\
&\leq K'_u\lambda_u + |u - v| && [\text{by induction}] \\
&\leq \left(\frac{K'_u}{\gamma} + 1\right)|u - v| && [\text{because } u \text{ did not yield to } v] \\
&\leq K'_I|u - v| && \left[\frac{K'_u}{\gamma} + 1 \leq K\right] \\
&= K'_I\lambda_v. && [\lambda_v = |u - v|]
\end{aligned}$$

□

Lemma 4.9.3. *Let q and q' be any two input points and let r be the distance between them. If $A = \mathbf{annulus}(q, 2r, \frac{6r}{\varepsilon^3})$ contains no input points, then q and q' are inside some cage contained in A for all intermediate meshes after each has been inserted.*

Sketch. Let $p_1, \ldots, p_k$ be all input points in $\mathbf{ball}(p, 2r)$ ordered by the order in which they were inserted. Clearly q and q' are among the p_i's. The proof is a straightforward induction on k, requiring us only to show that each insertion leaves the desired cage around the previous set. The constant $\frac{6}{\varepsilon^3}$ was carefully chosen to make this work. The full proof is Lemma 4.17.4 in the appendix. □

We can now prove that the output mesh has size linear in the input size.

Theorem 4.9.4. *If M is the output of the* NetMesh *algorithm for an input set P, then $|M| = O(|P|)$.*

Proof. Let Ω be any domain in the output. Let $p_1 \ldots, p_j$ be the vertices of P_Ω ordered such that for each $i = 3 \ldots j$, $\mathbf{f}_{P_i^\Omega}(p_i)/\mathbf{f}_{P_{i-1}^\Omega}(p_i) \geq \frac{12}{\varepsilon^3} + 1$, where $P_i = \{p_1, \ldots, p_i\}$. Lemma 4.9.3 guarantees that such an ordering can be found by a trivial greedy algorithm (see Lemma 4.17.3 for details of the construction).

In Chapter 3, we showed if P_Ω can be ordered this way then any well-spaced superset satisfying the bound in Lemma 4.9.2 has size $O(|P_\Omega|)$. So, in particular $|M_\Omega| = O(|P_\Omega|)$. Now, we observe that because every domain contains at least 2 input points, $\sum_\Omega |P_\Omega| < 2|P|$. So, the total mesh size can be bounded as $|M| \leq \sum_\Omega |M_\Omega| = O(\sum_\Omega |P_\Omega|) = O(|P|)$. □

4.10 Range Spaces and ϵ-Nets

In this section we discuss ideas and definitions from hypergraph and range space theory that we will need in our meshing algorithm. We will also give a distance measure derived from a range space that is useful for our analysis. A **range space** or **hypergraph** is a pair $(X, \mathcal{R})$ where X is a set and $\mathcal{R}$ is a collection of sets called **ranges**. A **range space ϵ-net** for $(X, \mathcal{R})$ is a subset N of X such that $N \cap R \neq \emptyset$ for all $R \in \mathcal{R}$ such that $|R \cap X| \geq \epsilon |X|$.

Throughout this discussion the ranges will be open balls in $\mathbb{R}^d$ including those with infinite radius, i.e. halfspaces. For a subset $M \subset \mathbb{R}^d$, define:

$$\mathcal{B}_M = \{B \ : \ B \text{ is a ball and } B \cap M = \emptyset\}.$$

A useful subset of $\mathcal{B}_M$ is the set of D-balls of M:

$$\mathcal{D}_M = \{B \in \mathcal{B}_M \ : \ B \text{ is a D-ball of } M\}.$$

The following geometric lemma is useful for translating between statements about D-balls and statements about arbitrary empty balls in the space.

Theorem 4.10.1 (D-ball Cover Theorem)**.** *If $M \subset \mathbb{R}^d$ and $B \in \mathcal{B}_M$ then B is covered by at most d D-balls of $\mathcal{D}_M$ and these d balls all share a common point.*

The proof is in Appendix 4.16.

Let $\mathcal{G}_{\mathcal{B}_M}$ be the graph with vertex set $\mathcal{B}_M$ and edges for each pair of balls that intersect. For any $x, y \in \mathbb{R}^d \setminus M$, let $\mathbf{d}_{\mathcal{B}_M}(x, y)$ be the length of the shortest path in $\mathcal{G}_{\mathcal{B}_M}$ between a ball containing x to a ball containing y. Define $\mathcal{G}_{\mathcal{D}_M}$ and $\mathbf{d}_{\mathcal{D}_M}$ similarly. These distances are related by the following lemma.

Lemma 4.10.2. *If $M \subset \mathbb{R}^d$ is finite then $\mathbf{d}_{\mathcal{D}_M} \leq 2\mathbf{d}_{\mathcal{B}_M}$*

Proof. Let $x, y \in \mathbb{R}^d$ be any pair of points and let $s = \mathbf{d}_{\mathcal{B}_M}(x, y)$. It will suffice to find D-balls $E_1, \ldots, E_{2s} \in \mathcal{D}_M$ such that $x \in E_1$, $y \in E_{2s}$, and each $E_i \cap E_{i+1}$ is nonempty. By the definition of $\mathbf{d}_{\mathcal{B}_M}$, there exists balls $B_1, \ldots, B_s \in \mathcal{B}_M$ such that $x \in B_1$, $y \in B_s$, and each $B_i \cap B_{i+1}$ is nonempty. Let z_i be a point in $B_i \cap B_{i+1}$ for $i = 1 \ldots s-1$ and define $z_0 := x$ and $z_s := y$. Now, by Theorem 4.10.1, there are d D-balls covering each B_i and they all have a common intersection. So, letting E_{2i-1} and E_{2i} be the D-balls among these that contain z_{i-1} and z_i gives the desired path of length at most $2s$ in $\mathcal{G}_{\mathcal{D}_M}$. □

4.11 Bounding the ply

In this section, we prove that the D-balls have constant ply, a fact that is useful for many parts of the analysis. In particular this is important for showing that the cost of a single insertion is constant. The constant ply can then be used to show that the degree of the intersection graph of the D-balls is bounded by a constant, and moreover, that the degree of every vertex of the Delaunay 1-skeleton is bounded by a constant. The latter bound implies that the total complexity of the Delaunay triangulation is linear in the number of vertices.

These results are known in the case of Voronoi diagrams with bounded aspect ratio [MTTW99], but we extend them to hold when there is a domain hierarchy rather than a single domain. To begin, we give some lemmas about the limited interaction between the different domains in the hierarchy.

Lemma 4.11.1. *Let (M, H) be a hierarchical mesh with cages. For any domain Ω, every D-ball B that intersects $B'_\Omega = \mathbf{ball}(c_\Omega, (1 - 2\delta - 2\gamma)r_\Omega)$ is contained in B_Ω.*

Proof. Suppose for contradiction that there exists a D-ball $B = \mathbf{ball}(c, r)$ such that both $B \cap B'_\Omega$ and $B \setminus B_\Omega$ are nonempty. Let z be the projection of c onto the sphere $\{x : |x - c| = (1 - \delta - \gamma)r_\Omega\}$. By the cage spacing properties, there is some vertex u in C_Ω such that $|u - z| < (\delta + \gamma)r_\Omega$. Next, $|c - z| \leq r - (\delta - \gamma)r_\Omega$ by our supposition and the triangle inequality. So, it follows that $|c - u| \leq |u - z| + |z - c| < r$. However, this implies that $u \in B$, contradicting the assumption that B is a D-ball. □

Lemma 4.11.2. *Let M be a τ-quality hierarchical mesh with parameters such that $\varepsilon + \varepsilon^2 < 1 - 2\delta - 2\gamma$. If Ω_1, Ω_2, Ω_3 are three nested domains in H_M, i.e. $\Omega_1 = \mathrm{p}(\Omega_2)$ and $\Omega_2 = \mathrm{p}(\Omega_3)$ then no D-ball intersects both Ω_1 and Ω_3.*

Proof. Suppose for contradiction that some ball B intersects both Ω_1 and Ω_3. Let c_2 and c_3 be the centers of Ω_2 and Ω_3 respectively. Since c_2 does not encroach the outside of Ω_3 and c_3 does not encroach the inside of Ω_3 we have that

$$\frac{1}{\varepsilon} r_{\Omega_3} \leq |c_2 - c_3| \leq \varepsilon r_{\Omega_2}. \tag{4.7}$$

Let x be a point of $B \cap \Omega_3$. By (4.7) and the triangle inequality, $|x - c_2| \leq (\varepsilon + \varepsilon^2) r_{\Omega_2}$. Since $\varepsilon + \varepsilon^2 < 1 - 2\delta - 2\gamma$, Lemma 4.11.1 implies that B is contained in B_{Ω_2} and therefore is disjoint from Ω_1, a contradiction. □

Lemma 4.11.3. *Let M be a τ-quality hierarchical mesh and let Ω be any domain in H_M. If B is a ball of radius r centered in B_Ω empty of points in M and $x \in B$, then*

$$\mathbf{f}_M^\Omega(x) \geq c_{4.11.3} r,$$

where $c_{4.11.3} = \frac{1}{24\tau^2}$

Proof. The proof of Lemma 6.1 from [HMP06] may be repeated verbatim here even though the Voronoi cells are defined differently because the proof only uses the the in-radius and out-radius conditions which are well-defined. □

Lemma 4.11.4. *Let M be a τ-quality mesh with $\varepsilon < \frac{1}{3}$. Let $B = \mathbf{ball}(c, r)$ be a D-ball corresponding to a simplex $\sigma \subset M$ and let x be a point of B. If Ω is a domain such that the ancestor of some vertex of σ in M_Ω is not the nearest neighbor of x in M_Ω, then*

$$\mathbf{f}_M^\Omega(x) \leq 3r.$$

Proof. Let u' be the ancestor of u in M_Ω assumed to exist, and thus $f_M^\Omega(x) \leq |x - u'|$. Since no domains are ε-encroached, $|u' - u| \leq \varepsilon |u' - v|$, and thus, by the triangle inequality $|u' - u| \leq \frac{\varepsilon |u - v|}{1 - \varepsilon}$. Since u and v are on the boundary of B, $|u - v| \leq 2r$ and so, using the assumption that $\varepsilon < \frac{1}{3}$, $|u - u'| \leq r$. Because $x \in B$, $|x - u| \leq 2r$. Using the triangle inequality and the preceding inequalities, we get $f_M^\Omega(x) \leq |x - u| + |u - u'| \leq 3r$. □

Theorem 4.11.5. *A τ-quality hierarchical mesh has ply at most $c_{4.11.5}$ where $c_{4.11.5}$ depends only on the meshing parameters.*

Proof. Let x be any point and let Ω be the domain containing x. Let S be the set of D-balls containing x. For any $\sigma \in \mathrm{Del}_M$, let B_σ be its D-ball and let Ω_σ be the least common ancestor domain of its vertices. Let n_x be the nearest neighbor of x in M_Ω and let Ω' be the child domain of Ω centered at n_x (if it exists).

There are two corner cases that we need to eliminate first: these are balls B_σ in S such that

1. $\Omega' = \Omega_\sigma$, or
2. $\Omega = \Omega_\sigma$ and B_σ is not centered in Ω.

In both of these cases, B_σ spans the bounding sphere of either Ω or Ω'. So, Lemmas 4.11.1 and 4.11.8 imply that there are only a constant number of such balls. Let S' be the subset of S formed by removing this constant sized set of balls.

Consider the following two subsets of S'.

$$S_1 = \{B_\sigma \in S' : \Omega = \Omega_\sigma \text{ or } \Omega = \mathrm{p}(\Omega_\sigma)\}$$
$$S_2 = \{B_\sigma \in S' : \mathrm{p}(\Omega) = \Omega_\sigma \text{ or } \mathrm{p}(\Omega) = \mathrm{p}(\Omega_\sigma)\}$$

By Lemma 4.11.1, S_1 and S_2 cover all of S'. Thus it will suffice to prove these sets have constant size. Ignoring the overlap, these two cases correspond to counting the balls that contain x that come from its own domain or children and those that come from a parent or sibling domain. In fact, the two cases are completely symmetric. The rest of the proof will show that $|S_1|$ is a constant. It can then be repeated for the second case by inserting S_2 and $\mathrm{p}(\Omega)$ in place of S_1 and Ω.

The set of vertices on simplices whose D-balls are in S_1 is $V = \{v \in \sigma : B_\sigma \in S_1\}$. Let v' denote the ancestor in M_Ω of any vertex $v \in V$, and define $V' = \{v' : v \in V\}$. We will first show that $|V| \leq c_{4.11.8}|V'|$. Then we will show that $|V'| \leq \alpha$, where $\alpha = \left(\frac{24\tau}{c_{4.11.3}}\right)^d$. Every D-ball in S_1 corresponds to some subset of $d+1$ points of V so will conclude that $|S_1| \leq \binom{|V|}{d+1} \leq (\alpha c_{4.11.8})^{d+1}$.

Claim: $|V| \leq c_{\mathbf{4.11.8}}|V'|$**.**
Fix some $v \in V'$. Let $U = \{u \in V : u' = v'\}$ It will suffice to show that $|U| \leq c_{4.11.8}$. If $v' \in \Omega$, then $v = v'$ and so $|U| = 1$. So, we may assume that v' is the center of some domain $\Omega_{v'}$ whose parent is Ω. Note that $B_\sigma \cap \Omega$ is nonempty for all $B_\sigma \in S_1$. So, any $u \in U$ came from a ball B_σ that spans the bounding sphere of $\Omega_{v'}$. Thus, Lemmas 4.11.1 and 4.11.8 implies that $|U| \leq c_{4.11.8}$.

Claim: $|V'| \leq \alpha$**, where** $\alpha = \left(\frac{24\tau}{c_{\mathbf{4.11.3}}}\right)^d$**.**
Let $r_{\min}$ and $r_{\max}$ be the minimum and maximum radii among the balls of S_1. By Lemma 4.11.3, $\mathbf{f}^\Omega_M(x) \geq c_{4.11.3} r_{\max}$. Lemma 4.11.4 implies that $\mathbf{f}^\Omega_M(x) \leq 3r_{\min}$. Combining these two facts, we see that all of the balls have radii that differ by at most a constant:

$$r_{\max} \leq \frac{3r_{\min}}{c_{4.11.3}}. \tag{4.8}$$

Consider any ball $B_\sigma \in S_1$ with a vertex $v \in \sigma$. The bounded aspect ratio condition implies that

$$f^\Omega_M(v') \geq \frac{r_{\min}}{\tau}. \tag{4.9}$$

The vertices $v' \in V'$ are not too far from x compared to the radii of the balls:

$$\begin{aligned}
|v' - x| &\leq |v - v'| + |v - x| && \text{[by the triangle inequality]} \\
&\leq |v - v'| + 2r_{\max} && [v, x \in B \in S_1] \\
&\leq \varepsilon \mathbf{f}^{\Omega}_M(v') + 2r_{\max} && \text{[non-encroachment]} \\
&\leq 3r_{\max}. && \left[\varepsilon \leq \frac{1}{3} \text{ and Lemma 4.11.4}\right] \qquad (4.10)
\end{aligned}$$

We can now show that $|V'| \leq \alpha$ by a volume packing argument. Specifically, let $U = \{\mathbf{ball}(v', \frac{r_{\min}}{2\tau}) : v' \in V'\}$. Note that $|U| = |V'|$. By (4.9), these balls are disjoint. By (4.10), these balls are contained in a ball of radius $4r_{\max}$. Applying (4.8), we conclude that $|U| \leq \left(\frac{24\tau}{c_{4.11.3}}\right)^d$. □

Corollary 4.11.6. *Every vertex v in a τ-quality hierarchical mesh M is in at most $(d+1)c_{4.11.5}$ Delaunay simplices.*

Proof. Let U be a set of $d+1$ points in $\mathbf{ball}(c, r)$ such that $v \in \text{conv}(U)$ and r is sufficiently small so that every D-ball with v on its boundary intersects U. By Theorem 4.11.5, there are only $c_{4.11.5}$ D-balls intersecting any point in U, so the total number of Delaunay simplices containing v is at most $(d+1)c_{4.11.5}$. □

Theorem 4.11.7. *Let M be a τ-quality mesh and let $\mathcal{G}_{\text{Del}}$ be the graph formed by the 1-skeleton of Del_M. The maximum degree of any node of $\mathcal{G}_{\text{Del}}$ is bounded by a constant that depends only on the meshing parameters.*

Proof. Let $v \in M$ be any vertex. Note that every simplex containing v in Del_M has a corresponding D-ball with v on its boundary. Corollary 4.11.6 implies that there are only $(d+1)c_{4.11.5}$ such D-balls. Each such ball contributes at most d edges, so the total is at most $d(d+1)c_{4.11.5}$. □

Lemma 4.11.8. *Let (M, H) be a hierarchical mesh with no encroached cages. Let $\Omega \in H$ be a domain such that $\mathbf{f}^{\Omega}_P(z) \leq K\mathbf{f}^{\Omega}_M(z)$ for all $z \in \Omega$. The number of vertices of M contained in $A = \mathbf{annulus}(c_\Omega, 2\varepsilon r_\Omega, r_\Omega)$ is at most some constant $c_{4.11.8}$ depending only on the meshing parameters.*

Proof. For points z in A, we know that $\mathbf{f}^{\Omega}_P(z) \geq \varepsilon r_\Omega$ because C_Ω is not encroached. It follows that for all $z \in M \cap A$, $\mathbf{f}^{\Omega}_M(z) \geq \frac{\varepsilon}{K} r_\Omega$. So, there must be disjoint balls of radius at least $\frac{\varepsilon}{2K} r_\Omega$ around each such z. Therefore, a simple packing completes the proof. □

Theorem 4.11.9. *Let M be a τ-quality mesh and let $\mathcal{G}_{\mathcal{D}}$ be the intersection graph of the D-balls of Del_M. The maximum degree of any node of $\mathcal{G}_{\mathcal{D}}$ is bounded by a constant $c_{4.11.9}$ that depends only on the meshing parameters.*

Proof. Let $B = \mathbf{ball}(c, r)$ be any D-ball and let Ω be the domain containing its center. Let S be the set of D-balls intersecting B. We will show that $|S| \leq c_{4.11.9}$ by bounding separately the number

of such balls with radius at least βr and those with radius less than βr, where β is a constant independent of the dimension.

Let $S_1 = \{b \in S : \text{radius}(b) \geq \beta r\}$. Observe that for any $b \in S_1$, $\mathbf{Vol}(b \cap \mathbf{ball}(c, 1+2\beta)) \geq \beta^d \mathbb{V}_d$. Theorem 4.11.5 implies that $\mathbf{ball}(c, r + 2\beta r)$ is covered at most t times by the D-balls of S_1 and thus by volume packing,

$$|S_1| \leq t(2 + \frac{1}{\beta})^d.$$

We will now bound the size of $S_2 = \{b \in S : \text{radius}(b) < \beta r\}$. Let $B' = \mathbf{ball}(c', r')$ be a D-ball of S_2 and let V be the vertices of the Delaunay simplex corresponding to B'. Let $h : \mathcal{D}_M \to M$ be a map that takes a D-ball B' to an arbitrary vertex its corresponding Delaunay simplex. Let $g : S \to M_\Omega$ be a map defined as $g(B') = \text{lca}_{M_\Omega}(h(B'))$. As a shorthand, we write $g(S)$ to denote $\bigcup_{B' \in S}\{g(B')\}$. The map g allows us to charge the balls of S to nearby vertices in M_Ω. In Lemma 4.11.10, we prove that

$$|g(S_2)| \leq c_{4.11.10}.$$

Then, in Lemma 4.11.11, we prove that

$$|g^{-1}(v)| \leq c_{4.11.11},$$

for all $v \in M_\Omega$. Together these allow us to conclude that

$$|S_2| = \sum_{v \in g(S_2)} |g^{-1}(v)| \leq c_{4.11.10} c_{4.11.11}.$$

So, setting $c_{4.11.9} = t(2 + \frac{1}{\beta})^d + c_{4.11.10} c_{4.11.11}$, we conclude that

$$|S| = |S_1| + |S_2| \leq c_{4.11.9}.$$

□

Lemma 4.11.10. *If S_2 is a collection of D-balls of radius at most βr intersecting a D-ball $B = \mathbf{ball}(c, r)$, then $|g(S_2)| \leq c_{4.11.10}$, where $\beta = \frac{c_{4.11.3}}{4} r$ and $c_{4.11.10}$ is a constant that depends only the meshing parameters.*

Proof. Fix some $B' \in S_2$ and let r' be its radius. Let Ω be the domain containing c. Let $u = h(B')$ and $v = g(B')$. If $u, v \in \Omega$ then $u = v$ and $|u - v| = 0$. The cage construction guarantees that if $u, v \notin \Omega$ then $|u - v| \leq (1 - \delta + \gamma)|c - v|$. Using the triangle inequality, we know that $|c - v| \leq r + 2r' + |u - v|$. Combining these inequalities, we get that

$$|c - v| \leq \alpha r, \tag{4.11}$$

where $\alpha = \frac{1+2\beta}{\delta - \gamma}$.

We want to prove that the vertices of $g(S_2)$ are not too close together. To do this, we will

bound the feature size at $v \in g(S_2)$. There are two cases to consider. First, if $|u - v| > \beta r$ then $\mathbf{f}_M^\Omega(v) \geq \beta r$ because the cage centered at v that contains u cannot intersect any point of M_Ω (other than v itself). Second, if $|u - v| \leq \beta r$ then $\mathbf{f}_M^\Omega(v) \geq \mathbf{f}_M^\Omega(z) - |z - u| - |u - v|$, where $z \in B \cap B'$. Since $z \in B$ and B is centered in Ω, Lemma 4.11.3 implies that $\mathbf{f}_M^\Omega \geq c_{4.11.3} r = 4\beta r$. Since z and u are in B', $|z - u| \leq 2r' \leq 2\beta r$. So in this case as well, we conclude that

$$\mathbf{f}_M^\Omega(v) \geq \beta r. \tag{4.12}$$

We can now complete the proof with a volume packing argument. For each $v \in g(S_2)$, we consider the ball $b_v = \mathbf{ball}(v, \frac{\beta}{2} r)$. Inequality (4.12) implies that these balls are disjoint. Moreover, (4.11) implies that these balls are all contained in $\mathbf{ball}(c, (\alpha + \frac{\beta}{2})r)$. It follows that the number of balls b_v can be at most $c_{4.11.10} = \left(\frac{2\alpha}{\beta} + 1\right)^d$. □

Lemma 4.11.11. *For all $v \in M_\Omega$, $|g^{-1}(v)| \leq c_{4.11.11}$, where $c_{4.11.11}$ is a constant that depends only on the meshing parameters.*

Proof. Let $U = \{h(B') : B' \in g^{-1}(v)\}$. Corollary 4.11.6 implies that $|g^{-1}(v)| \leq (d+1)c_{4.11.5}|U|$. So, it will suffice to prove that $|U| \leq \frac{c_{4.11.11}}{(d+1)c_{4.11.5}}$.

Fix $B' = \mathbf{ball}(c', r) \in g^{-1}(v)$ and let $u = h(B')$. If $u \in M_\Omega$ then $u = v$ and $|U| = 1$. So, we may assume that there is some domain $\Omega'' \in \text{children}(\Omega)$ centered at v. Let r'' be the radius of Ω''. There are two cases to consider.

Case 1: $u \in U_1 = \{u \in U : |u - v| \leq 2\varepsilon r''\}$**.** Lemma 4.11.1 implies that $B' \subset B_{\Omega''}$. Lemma 4.11.2 implies that $r' \geq (1 - 2\delta - 2\gamma - 2\varepsilon)r''$. Theorem 4.11.5 says that $B_{\Omega''}$ can only be covered $c_{4.11.5}$ times by the balls B'. So, by volume packing $|U_1| \leq \alpha$, where $\alpha = (1 - 2\delta - 2\gamma - 2\varepsilon)^{-d}$.

Case 2: $u \in U_2 = \{u \in U : |u - v| > 2\varepsilon r''\}$**.** Lemma 4.11.8 implies that $|U_2| \leq c_{4.11.8}$.

We now conclude that $|U| = |U_1| + |U_2| \leq \alpha + c_{4.11.8}$. Choosing $c_{4.11.11} = (d+1)c_{4.11.5}(\alpha + c_{4.11.8})$ completes the proof. □

4.12 The Quality Invariant

In this section, we will prove that NetMesh maintains the quality invariant throughout the course of the algorithm. As shown in Section 4.11, this is an important property to have. Our goal will be to prove the following.

Theorem 4.12.1. *For any input, the intermediate meshes of the* NetMesh *algorithm are τ''-quality, where τ'' depends only on the mesh parameters.*

The proof will follow directly from Lemmas 4.12.2 and 4.12.8 below. The former guarantees that the quality is bounded after every call to InsertInput. The latter guarantees that the quality is bounded throughout the Clean operation.

4.12.1 Quality during input insertion.

In this section, we will show that starting with a τ-quality mesh and inserting an input point results in a τ'-quality mesh where τ' is a constant that depends only on the meshing parameters. Throughout this section, let M be the τ-quality starting mesh and let M' be the mesh after executing InsertInput(v) for some $v \in P$. Whenever we refer to a cell $\text{Vor}_M(u)$ or $\text{Vor}_{M'}(u)$, r_u and r'_u denote the respective in-radii and R_u, R'_u denote the out-radii. We say that a cell $\text{Vor}_M(u)$ is **caged** during InsertInput if a new cage is added that is contained in $\text{Vor}_M(u)$.

Lemma 4.12.2. *There is a constant τ' depending only on the meshing parameters such that the mesh after every call to* InsertInput *is τ'-quality.*

Proof. The Clean procedure explicitly guarantees that the starting mesh M is τ-quality. We need to prove that all cells $\text{Vor}_{M'}(u)$ (excepting cage vertices) have aspect ratio at most τ', i.e. $\frac{R'_u}{r'_u} \leq \tau'$. Fix one such u. There are four different cases to consider:

1. The Voronoi cell in M, $\text{Vor}_M(u)$, had aspect ratio at most τ.
2. $u = C_\Omega$ is a newly created cage.
3. u was a cage vertex in M that got released.
4. $u = v$ is the newly inserted input vertex.

Case 1: $\text{Vor}_M(u)$ had aspect ratio at most τ. If $\text{Vor}_M(u)$ was caged during InsertInput then Lemma 4.12.4 implies that $\text{Vor}_{M'}(u)$ has aspect ratio at most $c_{4.12.4}$. Otherwise, Lemma 4.12.3 implies that $r_u \leq c_{4.12.3} r'_u$. In this case, the out-radius cannot go up with the addition of more points so $R'_u \leq R_u$. Thus, since $R_u \leq \tau r_u$, we get that $R'_u \leq c_{4.12.3} \tau r'_u$ as desired.
Case 2: $u = C_\Omega$ is a newly created cage. For this case, Lemma 4.12.7 implies that $R'_u \leq c_{4.12.7} r'_u$.
Case 3: u was a cage vertex in M that got released. If u is caged then Lemma 4.12.4 implies that its aspect ratio is at most $c_{4.12.4}$. Otherwise, Lemma 4.12.3 implies that $r_u \leq c_{4.12.3} r'_u$. Let C_Ω be the released cage and let c be its center. The cage spacing guarantees that $r_\Omega \leq s r_u$, where $s = (\delta - \gamma)(1 - \delta - \gamma)$. So, it follows that

$$r_\Omega \leq c_{4.12.3} s r'_u. \tag{4.13}$$

Now, we must consider the cases where v encroached the inside or the outside of Ω. For an outer encroachment, Lemma 4.12.5 implies that $R_c \leq c_{4.12.5} r_\Omega$, where R_c is the out-radius of $\text{Vor}_M(C_\Omega)$. The out-radius of a cage is strictly greater than the out-radius of its cage vertices, so $R'_u < R_c$ and thus by (4.13), $R'_u < c_{4.12.5} c_{4.12.3} s r'_u$.

For an inner encroachment, we call GrowCage, which conditionally adds a larger cage around the existing cage before releasing it. Lemma 4.12.6 implies that $R'_u \leq c_{4.12.6} r_\Omega$. So, (4.13) implies $R'_u \leq c_{4.12.6} c_{4.12.3} s r'_u$.
Case 4: $u = v$ is the newly inserted input vertex. Note that in all of the preceding cases, inserting u only increased the aspect ratio bound. Consequently, by the time InsertInput actually

adds u to the mesh, it is added to a quality mesh M'' with the same domain hierarchy as M'. Let Ω be the domain u is inserted into. Let w be the nearest vertex to u in M''_Ω. Note that u is ε-medial in M'', for otherwise we would have created a new domain or yielded when inserting w, and so $\mathbf{f}^\Omega_{M''}(u) \leq \frac{1}{\varepsilon}|u-w| \leq \frac{3}{\varepsilon}r'_u$. Since M'' is quality and the D-ball centered at the farthest corner of $\mathrm{Vor}_{M'}(u)$ is empty in M'', Lemma 4.11.3 implies that $R'_u \leq c_{4.11.3}\mathbf{f}^\Omega_{M''}(u)$. Thus, $R'_u \leq \frac{3c_{4.11.3}}{\varepsilon}r'_u$. □

Lemma 4.12.3. *If* $\mathrm{Vor}_{M'}(u)$ *is cell that is not caged during* InsertInput(v)*, then* $r_u \leq c_{4.12.3}r'_u$.

Proof. We show the nearest neighbor of $\mathrm{Vor}_{M'}(u)$ cannot be too close. If it has a new nearest neighbor, it can only be from a neighboring cage recently added or the new input vertex. Since u does not encroach any new cages, they can only decrease the in-radius by a $1-\varepsilon$ factor. Since it was not caged, v must have been medial and therefore r'_u can only go down by a $\frac{\varepsilon}{2}$ factor. So, choosing $c_{4.12.3} = \frac{2}{\varepsilon(1-e)}$ suffices to yield $r_u \leq c_{4.12.3}r'_u$ as desired. □

Lemma 4.12.4. *If* $\mathrm{Vor}_{M'}(u)$ *is caged during* InsertInput(v)*, then* $R'_u \leq c_{4.12.4}r'_u$*, where* $c_{4.12.4}$ *depends only on the meshing parameters.*

Proof. Let C_Ω be the cage. The new Voronoi cell is contained in the newly formed domain, so $R'_u \leq r_\Omega$. If any other point was added to Ω, then it must have been v and so by construction, $r'_u \geq \frac{\varepsilon}{2}r_\Omega$. So, it suffices to choose $c_{4.12.4} = \frac{2}{\varepsilon}$. □

The following two lemmas show that cage vertices released in the algorithm have out-radii bounded by a constant times the radius of the cage they belonged to.

Lemma 4.12.5. *Let* C_Ω *be a cage in a hierarchical mesh* M *and let* $\Omega' = \mathrm{p}(\Omega)$*. Suppose there exists a point* v *that is* ε*-medial in* Ω' *and outer encroaches* C_Ω*. If* $\mathrm{Vor}_M(C_\Omega)$ *has aspect ratio at most* τ *then the out-radius of* $\mathrm{Vor}_M(C_\Omega)$ *is at most* $c_{4.12.5}r_\Omega$*, where* $c_{4.12.5}$ *depends only on the meshing parameters.*

Proof. Let R denote the out-radius of $\mathrm{Vor}_M(C_\Omega)$ and let c be the center of C_Ω. Since the aspect ratio is at most τ, it follows that $R \leq \tau\mathbf{f}^{\Omega'}_M(c)$. Let w be the second nearest neighbor of v in $M_{\Omega'}$. Then, we can bound $\mathbf{f}^{\Omega'}_M(c)$ as follows.

$$\begin{aligned}
\mathbf{f}^{\Omega'}_M(c) &\leq \mathbf{f}^{\Omega'}_M(v) + |c-v| && \left[\mathbf{f}^{\Omega'}_M \text{ is 1-Lipschitz}\right] && (4.14)\\
&\leq |v-w| + |c-v| && [\text{by the choice of } w] && (4.15)\\
&\leq (1+\frac{1}{\varepsilon})|c-v| && [v \text{ is } \varepsilon\text{-medial}] && (4.16)\\
&\leq \frac{1+\varepsilon}{\varepsilon^2}r_\Omega. && [v \text{ encroaches } \Omega] && (4.17)
\end{aligned}$$

So, it suffices to choose $c_{4.12.5} = \frac{\tau(1+\varepsilon)}{\varepsilon^2}$. □

Lemma 4.12.6. *Let* C_Ω *be a cage in a hierarchical mesh* M*. Let* M' *be the resulting mesh after* GrowCage(C_Ω)*. If* $\mathrm{Vor}_M(C_\Omega)$ *has aspect ratio at most* τ *then the out-radius of* $\mathrm{Vor}_{M'}(u)$ *is at most* $c_{4.12.6}r_\Omega$ *for all* $u \in C_\Omega$*, where* $c_{4.12.6}$ *depends only on the meshing parameters.*

Proof. Let R_u denote the out-radius of $\mathrm{Vor}_{M'}(u)$. In the GrowCage routine, either a new cage is added or it is not. In the former case, the new cage has radius $\frac{r_\Omega}{\varepsilon}$ so $R_u \leq \frac{r_\Omega}{\varepsilon}$ in this case. If the new cage is not added, it is because, $r_C \leq \frac{r_\Omega}{\varepsilon}$. Because we assumed that the Voronoi cell of C_Ω had aspect ratio at most τ, it follows that $R_C \leq \tau r_C \leq \frac{\tau r_\Omega}{\varepsilon}$. By the definition of R_C, we have that $R_u \leq R_C$ and thus $R_u \leq \frac{\tau r_\Omega}{\varepsilon}$ as desired. □

Lemma 4.12.7. *If $u = C_\Omega$ is a cage added during* InsertInput*, then the $R'_u \leq c_{4.12.7} r'_u$, where $c_{4.12.7}$ depends only on the meshing parameters.*

Proof. If u is caged, then Lemma 4.12.4 implies $\mathbb{R}'_u \leq c_{4.12.4} r'_u$. So, we may assume that u is not caged. Let c be the center of Ω and Ω' is the previous domain that C_Ω as inserted into. So, $\mathbf{f}_M^{\Omega'}(c) = 3r_u$ and thus Lemma 4.12.3 implies that

$$\mathbf{f}_M^{\Omega'}(c) \leq 3c_{4.12.3} r'_u. \tag{4.18}$$

Let B be the D-ball centered at the far corner x of $\mathrm{Vor}_{M'}(u)$ and let r be its radius. So, $R'_u \leq r + r_\Omega$. Since C_Ω is not encroached, we have that $r_\Omega \leq r$ and thus

$$R'_u \leq 2r. \tag{4.19}$$

Since B was empty in M, Lemma 4.11.3 implies that there is a $y \in \Omega$ such that

$$\mathbf{f}_M^{\Omega'}(y) \geq c_{4.11.3} r. \tag{4.20}$$

So, $\mathbf{f}_M^{\Omega'}(y) \leq \mathbf{f}_M^{\Omega'}(c) + r_\Omega$ by the Lipschitz property of $\mathbf{f}_M^{\Omega'}$. Next, $r_\Omega < r'_u$, so (4.18) implies that $f_M^{\Omega'}(y) \leq (3c_{4.12.3} + 1) r'_u$. So, by (4.20),

$$r \leq c_{4.11.3}(3c_{4.12.3} + 1) r'_u. \tag{4.21}$$

Therefore, (4.19) and (4.21) imply that $R'_u \leq 2c_{4.11.3}(3c_{4.12.3}+1) r'_u$. Choosing $c_{4.12.7} = 2c_{4.11.3}(3c_{4.12.3}+1)$ completes the proof. □

4.12.2 Quality during the refinement process.

Lemma 4.12.8 (Clean preserves quality)**.** *Let M' be any intermediate mesh in the course of running* Clean(M) *on a τ'-quality mesh M. Then, M' is τ''-quality, where τ'' depends only on the meshing parameters.*

Proof. Let $V = \mathrm{Vor}^{\Omega}_{M'}(v)$ for some $v \in M'_\Omega$ and $\Omega \in H_{M'}$. We need to prove that the aspect ratio $\frac{R_V}{r_V}$ is at most τ''. There are two cases: V is the Voronoi cell of a vertex or V is the Voronoi cell of a cage.

Case 1: V is the Voronoi cell of the vertex v. First, we observe that there is D-ball $B = \mathbf{ball}(x, R_V)$ centered on the farthest corner of V. By Lemma 4.11.3, $R_V \leq 24\tau'^2 \mathbf{f}_M^{\Omega}(v)$. Now,

we observe that $r_V \geq \frac{1}{3}\mathbf{f}^{\Omega}_{M'}(v)$ by Lemma 4.6.2. We apply Lemma 4.18.1 to get that $r_V \geq \frac{1}{3K'}\mathbf{f}^{\Omega}_{M}(v)$. So, we get that $\frac{R_V}{r_V} \leq 72K'\tau'^2$.

Case 2: V is the Voronoi cell of a cage centered at v. Since no new cages are created during CLEAN, it must be that $v \in M$. Let c be the center of the cage C that contains v. The cage spacing guarantees that $r_V \geq \frac{\delta-2\gamma}{2}|c-v|$. Let s be the Steiner point whose insertion caused C to be released. Let M'' be the mesh it was inserted into and let Ω' be the domain it encroached. Now, R_V cannot be larger than the out-radius of $\mathrm{Vor}(C)$ in M'', which, in turn, is at most R_C, the out-radius of $\mathrm{Vor}(C)$ in M because Voronoi cells can only shrink during cleaning. We now bound R_C as follows:

$$\begin{aligned} R_C &\leq \tau' r_C && \big[\text{because } M \text{ is } \tau'\text{-quality}\big] && (4.22)\\ &\leq \tau' \mathbf{f}^{\Omega'}_{M}(c) && [\text{by Lemma 4.6.2}] && (4.23)\\ &\leq \tau' K' \mathbf{f}^{\Omega'}_{M''}(c). && [\text{by Lemma 4.18.1}] && (4.24) \end{aligned}$$

Now, since s was inserted by a CLEAN operation into Ω', $\mathbf{f}^{\Omega'}_{M''}(c) \leq 3|c-s|$. Moreover, since s encroaches, we have that $|c-s| \leq \frac{|c-v|}{\varepsilon}$. So, we conclude that

$$R_V \leq \frac{3\tau' K'(1+\varepsilon)}{\varepsilon(1-\varepsilon)(\delta-2\gamma)} r_v.$$

□

4.13 Point Location Analysis

Definition. *A vertex $v \in M$ **touches** an uninserted point $u \in P \setminus M$ if when v was inserted into M there were intersecting D-balls B_u and B_v containing u and v respectively.*

The quality invariant and Theorem 4.8.1 guarantee that only a constant number of balls are created or destroyed during an insertion, so the total amount of point location work done on any input point is $O(t)$, where t is the number of times it was touched.

Theorem 4.13.1. *The total cost of point location in the* NETMESH *algorithm is $O(n \log n)$.*

Proof. As noted before, it suffices to count the number of touches on uninserted input points throughout the algorithm. Since there are only $O(\log n)$ rounds, it will suffice to show that no input point can be touched more than a constant number of times in a single round.

Let M be the mesh at the start of a round. Consider any point $p \in P$. We will show that p cannot be touched more than a constant number of times in this round. By definition, a point x touches p if $\mathbf{d}_{\mathcal{D}_{M'}}(p,x) \leq 1$ in the mesh M' just prior to inserting x. So, it follows that $\mathbf{d}_{\mathcal{B}_M}(p,x) \leq 1$ because D-balls in M' are empty of points of M. Moreover, by Lemma 4.10.2, $\mathbf{d}_{\mathcal{D}_M}(p,x) \leq 2$. Therefore, the set of points that can touch p this round are all contained in one of the constant number of D-balls that are within 2 hops of p in $\mathcal{G}_{\mathcal{D}_M}$. In Lemma 4.13.2 below, we show that only a constant

number of points are added to any D-ball in a single round. Thus, the total number of points that can touch p in a round is at most a constant. □

Lemma 4.13.2. *In any round starting with a mesh M, at most a constant number of points are added to any D-ball of M.*

Proof. Fix a particular round and let B be a D-ball of M. Let M' denote the mesh at the end of the round. Let P' denote the input points of M'. We wish to upper bound the number of points of M' in B. By standard mesh size analysis,

$$\begin{aligned} |M' \cap B| &= \sum_{\Omega \in H_{M'}} O\left(\int_{B_\Omega \cap B} \frac{dx}{\mathbf{f}^\Omega_{P'_\Omega}(x)^d} \right) \\ &\leq \sum_{\Omega \in H_{M'}} O\left(\int_{B_\Omega \cap B} \frac{dx}{\mathbf{f}^\Omega_{P'_\Omega \cup M_\Omega}(x)^d} \right). \end{aligned} \tag{4.25}$$

Lemma 4.13.3 shows that there are only $c_{4.13.3}$ terms in this summation and Lemma 4.13.6 shows each term is at most $c_{4.13.6}$. Thus, we conclude that $|M' \cap B| \leq c_{4.13.3} c_{4.13.6}$. □

Lemma 4.13.3. *Let M and M' be the meshes before and after a round of the* NetMesh *algorithm. For any D-ball B in M', at most a constant number of domains of $H_{M'}$ intersect B.*

Proof. Let x be the center of B. There are only a constant number of domains of H_M intersecting B, because each contains a D-ball intersecting B and Theorem 4.8.1 implies there can only be a constant number of such balls. Any newly created domains must have been caused by the insertion of an input point $y \in M' \setminus M$. However, if the new domain intersects B then either y caused a cage from M to grow or $\mathbf{d}_{\mathcal{B}_M}(x, y) \leq 5$. In either case, there are only a constant number of new domains intersecting B. □

In the following lemmas, we fix a particular domain Ω and use the following simplified notations. We number the k input points added this round as $P'_\Omega \setminus P_\Omega = \{p_1, \ldots, p_k\}$ where the ordering is the one given in Lemma 4.13.4 below. The part of the ball of Ω contained in B is defined as $A = B_\Omega \cap B$. The **near input points** are denoted Q and are formally defined as

$$Q = \{q \in P'_\Omega \setminus P_\Omega : \mathbf{d}_{\mathcal{B}_M}(q, x) \leq 5 \text{ for some } x \in B\}.$$

The index set of the **far input points** is $I = \{i : p_i \notin Q\}$ and $I_0 = I \cup \{0\}$. Let $S_i = M_\Omega \cup \{p_1, \ldots, p_i\}$ and set $S_0 = M$. The function f_i is equal to $\mathbf{f}^\Omega_{S_i}$.

We partition the set A into pieces based on which far point was the last to affect the feature size:

$$U_j = \{x \in A : \max\{i \in I : f_i(x) \neq f_{i-1}(x)\} = j\}$$

and

$$U_0 = A \setminus \bigcup_{j \in I} U_j.$$

Define h_i and h_{ij} as

$$h_i = \int_A \frac{dx}{f_i(x)^d}$$

and

$$h_{ij} = \int_{U_j} \frac{dx}{f_i(x)^d}.$$

Since A is the disjoint union of $\{U_j : j \in I_0\}$,

$$h_i = \sum_{j \in I_0} h_{ij}.$$

Lemma 4.13.4. *There exists an ordering $\{p_1, \ldots, p_k\}$ of $P'_\Omega \setminus P_\Omega$, such that $h_i - h_{i-1} \leq c_{4.13.4}$, for all $i = 1 \ldots k$.*

Proof. The desired ordering is a so-called **well-paced** ordering. It is one for which $\mathbf{f}^\Omega_{S_i}(p_i) \geq \alpha \mathbf{f}^\Omega_{S_{i-1}}(p_i)$ for all i. In previous work [MPS08], we showed that the change in the feature size integral over any domain is at most a constant after inserting a well-paced point. Calling this constant $c_{4.13.4}$, it will suffice to show that $P'_\Omega \setminus P_\Omega$ is well-paced with respect to M_Ω. This requires the same case analysis as used in Lemma 4.17.4, though it is easier in this case because we only require the inputs to be well-paced with respect to $M - \Omega$ rather than the stronger condition that they be well-paced with respect to the bounding cage. Alternatively, one could get an explicit ordering directly from the algorithm by keeping a list for each domain, appending new input points to the list corresponding to the domain that contains them, and appending the list for a released domain to the list of its parent. That the resulting list satisfies the well-paced condition is immediate from the algorithm and the feature size invariant. □

Lemma 4.13.5. *$|Q| \leq c_{4.13.5}$, where $c_{4.13.5}$ is a constant that depends only on d and the meshing parameters.*

Proof. By definition, $q \in Q$ implies that $\mathbf{d}_{\mathcal{B}_M}(q, x) \leq 5$ for some $x \in B$. It is easily checked that Theorem 4.11.9 and Lemma 4.10.2 implies that at most $10c_{4.11.9}$ D-balls can contain points of Q. Each round is defined by selecting only a constant size net from each D-ball. So each D-ball only contributes at most a constant number of points to Q and thus the total size of Q is at most a constant. □

Lemma 4.13.6. *$\int_{B_\Omega \cap B} \frac{dx}{\mathbf{f}^\Omega_{P'_\Omega \cup M_\Omega}(x)^d} \leq c_{4.13.6}$, where $c_{4.13.6}$ is a constant that depends only on d and the meshing parameters.*

Proof. In our simplified notation, the statement reduces to proving that $h_k \leq c_{4.13.6}$. Writing h_k as

a telescoping sum, we get

$$\begin{aligned} h_k &= h_0 + \sum_{i=1}^{k} (h_i - h_{i-1}) \\ &= h_0 + \sum_{p_i \in Q} (h_i - h_{i-1}) + \sum_{i \in I} (h_i - h_{i-1}) \\ &\leq h_0 + c_{4.13.4} c_{4.13.5} + \sum_{i \in I} (h_i - h_{i-1}) \qquad \text{[by Lemmas 4.13.4 and 4.13.5]} \end{aligned}$$

Far input points cannot change the feature size by very much. This is formalized in Lemma 4.19.2, where it is proven that

$$\sum_{i \in I} (h_i - h_{i-1}) \leq (c_{4.19.2} - 1) h_0.$$

The feature size in an empty ball cannot be too small compared to its radius. Specifically, Lemma 4.19.3 shows that $h_0 \leq c_{4.19.3}$. So setting $c_{4.13.6} = c_{4.13.4} c_{4.13.5} + c_{4.19.2} c_{4.19.3}$ suffices to complete the proof. See Appendix 4.19 for Lemmas 4.19.2 and 4.19.3 and their proofs. □

4.14 Finishing the mesh

The finishing process takes a hierarchical quality mesh and returns a well-spaced mesh.

Theorem 4.14.1. *Given a hierarchical quality mesh, the* FinishMesh *procedure runs in $O(m)$ time, where m is the size of the output mesh.*

Proof. The GrowCage and Clean procedures preserve the quality of the mesh, so each insertion takes constant time. There is no point location work to be done, so the total running time is linear in the number of points added. □

4.15 Conclusion and Future Work

In this chapter, we have given an algorithm for generating quality hierarchical meshes of point sets with size $O(n)$ in $O(n \log n)$ time. We also showed how to extend these hierarchical meshes to traditional well-spaced meshes in optimal output-sensitive time $O(n \log n + m)$. The algorithm and its analysis introduce novel uses of ϵ-nets and the linear-size meshing theory introduced in [MPS08].

Future Work We have restricted our discussion to the case of point set inputs. We expect it should now be possible to design a work efficient algorithm for inputs with higher dimensional features such as segments and faces. The algorithm presented is basically a work efficient parallel algorithm. It should be possible to show the present algorithm runs in polylog parallel time with no increase in work and thus beating the time and work bounds in parallel SVR [HMP07].

Yet another issue is integrating ideas from the NetMesh algorithm into the already relatively fast SVR code [AHMP07]. Future experiments in this direction are in order. The algorithm removes the spread term in the run time for the mesh based persistent homology algorithms [HMOS10]. It may also have applications for efficient surface reconstruction especially in the higher dimensional cases[Dey07].

4.16 The Ball Cover Theorem

In this section we prove that any open ball that contains no points from a set P is covered by d D-balls of Del_P. The main tool we use is Carathéodory's Theorem.

4.16.1 Carathéodory's Theorem

Carathéodory's Theorem is a classic result on convex sets that is critical to our proof of the D-ball Cover Theorem (Theorem 4.10.1).

Theorem 4.16.1 (Carathéodory's Theorem)**.** *Let $A \subset \mathbb{R}^d$. If $x \in \mathrm{cone}(A)$ then $x \in \mathrm{cone}(A')$ for some $A' \subseteq A$ such that $|A'| \leq d$. If $x \in \mathrm{conv}(A)$ then $x \in \mathrm{conv}(A')$ for some $A' \subseteq A$ such that $|A'| \leq d+1$.*

We will need the following extended form of Carathéodory's Theorem for $\mathcal{V}$-polyhedra, i.e. those formed by the Minkowski sum of a polytope and a cone.

Corollary 4.16.2. *If $x \in \mathrm{conv}(A) + \mathrm{cone}(B)$, then there exist subsets $A' \subset A$ and $B' \subset B$ such that*

$$x \in \mathrm{conv}(A') + \mathrm{cone}(B') \text{ and } |A'| + |B'| \leq d+1.$$

Moreover, if $|A'| = 0$ then $|B'| \leq d$.

Proof. Using the cone form of Carathéodory's Theorem, it suffices to observe that

$$x \in \mathrm{conv}(A) + \mathrm{cone}(B) \text{ if and only if } \begin{bmatrix} x \\ 1 \end{bmatrix} \in \mathrm{cone} \begin{bmatrix} A & B \\ 1 & 0 \end{bmatrix}.$$

□

We will also make use of the following technical lemma related to $\mathcal{V}$-polyhedra.

Lemma 4.16.3. *If $\mathrm{cone}(v) \subseteq \mathrm{conv}(A) + \mathrm{cone}(B)$ for some $v \in \mathbb{R}^d$ then $v \in \mathrm{cone}(B)$.*

Proof. We will prove the contrapositive. Suppose that $v \notin \mathrm{cone}(B)$. Then for a sufficiently large t, $\mathbf{d}(tv, \mathrm{cone}(B)) > \max_{a \in A} |a|$. Let $z = a + b$ be the nearest point of $\mathrm{conv}(A) + \mathrm{cone}(B)$ to tv, where

$a \in \text{conv}(A)$ and $b \in \text{conv}(B)$.

$$\begin{aligned} |tv - z| &\geq |tv - b| - |z - b| \\ &= |tv - b| - |a| \\ &> 0. \end{aligned}$$

Thus, $tv \notin \text{conv}(A) + \text{cone}(B)$ and therefore $\text{cone}(v) \not\subseteq \text{conv}(A) + \text{cone}(B)$. □

Lemma 4.16.4. *If c is a point in the $\mathcal{V}$-polyhedron $\text{conv}(Q) + \text{cone}(H)$, where $c \neq 0$ and $Q \neq \emptyset$, then for all $x \in \mathbb{R}^d$, either or both of the following hold:*

1. *$x^T(q - c) \geq 0$ for some $q \in Q$, or*
2. *$x^T h > 0$ for some $h \in H$.*

Proof. Let the coefficients α_i and β_j be such that

$$c = \sum_{i=1}^{|Q|} \alpha_i q_i + \sum_{j=1}^{|H|} \beta_j h_j,$$

where $\alpha_i, \beta_j \geq 0$ and $\sum \alpha_i = 1$. Suppose for contradiction that $x^T(q_i - c) < 0$ for all $q_i \in Q$, and $x^T h_j \leq 0$ for all $h_j \in H$. So, it follows that

$$\sum_{i=1}^{|Q|} \alpha_i x^T(q_i - c) + \sum_{j=1}^{|H|} \beta_j x^T h_j < 0.$$

Factoring this expression implies that

$$x^T \left(\sum_{i=1}^{|Q|} \alpha_i q_i + \sum_{j=1}^{|H|} \beta_j h_j - \sum_{i=1}^{|Q|} \alpha_i c \right) < 0.$$

However, the left side of the above inequality simplifies to $x^T(c - c) = 0$, a contradiction. □

Lemma 4.16.5. *If $c \in \text{cone}(H)$ for some $H \subset \mathbb{R}^d$ and B_c is a ball centered at c that does not contain the origin, then for all $x \in B_c$, there exists $h \in H$ such that $x^T h > 0$.*

Proof. Since $c \in \text{cone}(H)$, there are nonnegative coefficients $\{\beta_h\}_{h \in H}$ such that $c = \sum_{h \in H} \alpha_h h$. Fix any $x \in B_c$. Since $0 \notin B_c$, $|c - x| < |c - 0|$ and therefore $x^T c > \frac{|x|}{2} > 0$. Suppose for contradiction that $x^T h \leq 0$ for all $h \in H$. Then $x^T c = \sum_{h \in H} \alpha_h x^T h \leq 0$, a contradiction. □

4.16.2 A simpler version of the D-ball Cover Theorem

We warm up with a simpler version of the main result. It deals with the special case of balls centered in bounded Voronoi cells. It only proves the weaker bound of $d + 1$ rather d balls to cover and it does not prove that the covering balls are pairwise intersecting.

We start with a lemma that gives a sufficient condition for a point x to be in a bounded D-ball B_q.

Lemma 4.16.6. *Let c, q, and v be three points such that $v^T q = v^T c = 0$. Let B_q be an open ball centered at q with v on its boundary. Let B_c be an open ball centered at c that does not contain v. If $x \in B_c$ and $x^T(q-c) \geq 0$ then $x \in B_q$.*

Proof. Let $x \in B_c$ and $x^T(q-c) \geq 0$ according to the hypothesis. Since $x^T(q-c) \geq 0$, we have that

$$x^T q \leq x^T c. \tag{4.26}$$

Since $x \in B_c$ and $v \notin B_c$, $|c-x| < |c-v|$ and therefore because $v^T c = 0$,

$$x^T x - 2x^T c < v^T v. \tag{4.27}$$

It will suffice to prove that $|q-x| < |q-v|$. This follows from the following inequalities.

$$\begin{aligned} |q-x|^2 &= q^T q - 2x^T q + q^T q \\ &\leq q^T q - 2x^T c + q^T q && \text{[by (4.26)]} \\ &\leq v^T v + q^T q && \text{[by (4.27)]} \\ &\leq |q-v|^2. && \big[\text{because } v^T q = 0\big] \end{aligned}$$

□

Theorem 4.16.7. *Let M be a finite set. For any ball $B_c \in \mathcal{B}_M$ centered in a bounded Voronoi cell, there is a collection of at most $d+1$ D-balls that cover B_c.*

Proof. Let c be the center of B_c and let v be the nearest neighbor of c in M. By Carathéodory's Theorem, there exists a subset Q of the corners of $\mathrm{Vor}(v)$ such that $c \in \mathrm{conv}(Q)$ and $|Q| \leq d+1$. So, for any $x \in \mathbb{R}^d$, there is a $q \in Q$ such that $x^T(q-c) \geq 0$. Each $q \in Q$ corresponds to a D-ball B_q of radius $|q-v|$. Without loss of generality, we may assume v is the origin. Therefore, Lemma 4.16.6 implies that $x \in B_q$. □

4.16.3 The D-ball Cover Theorem

To prove the more general ball cover theorem, we need to be more careful to deal with infinite D-balls. The infinite D-balls are those corresponding to the facets of the convex hull of M. Thus, they are in correspondence with the unbounded 1-faces of Vor_M. The Voronoi cells of Vor_M are $\mathcal{V}$-polyhedra. That is, they can be written as the Minkowski sum of a convex polytope and a polytopal cone. So, there exists finite sets A, B such that $\mathrm{Vor}_M(v) = \mathrm{conv}(A) + \mathrm{cone}(B)$. Moreover, the set A is a subset of the Voronoi corners and B is a subset of the normals of the facets of $\mathrm{conv}(M)$. Thus, the points of A and the vectors of B are all in correspondence with the D-balls of $\mathcal{D}_M$.

Theorem 4.16.8 (The D-ball Cover Theorem)**.** *For all $B \in \mathcal{B}_M$, there exist d D-balls $b_1, \ldots, b_d \in \mathcal{D}_M$ such that $B \subseteq \bigcup_{i=1}^{d} b_i$. Moreover, these D-balls have a nonempty, common intersection.*

Proof. Let c be the center of B and let v be the nearest neighbor of c in M. Consider the ray starting from v that passes through c parameterized as $r(t) = v + t(c - v)$ for $t \geq 0$. We must distinguish between the cases where $r(t) \in \text{Vor}(v)$ for all $t \geq 0$ and where there is some t for which $r(t) \notin \text{Vor}(v)$.

Case 1: $r(t) \in \text{Vor}(v)$ for all $t \geq 0$. Voronoi cells are $\mathcal{V}$-polyhedra and can be written as $\text{Vor}(v) = \text{conv}(Q) + \text{cone}(H)$ where Q is a subset of D-ball centers and H is a set of normals of unbounded D-balls. Without loss of generality, we may assume that $v = 0$. By Carathéodory's Theorem, there is a subset $H' \subseteq H$ of size d such that $c \in \text{cone}(H')$. So, Lemma 4.16.5 implies that for all $x \in B$ there is an $h \in H'$ for which $x^T h > 0$. So, B is covered by the d unbounded D-balls corresponding to the vectors in H'. Moreover, $c^T h > 0$ and therefore, the chosen D-balls have a common intersection at c.

Case 2: $r(t) \notin \text{Vor}(v)$ for some $t \geq 0$. In this case, the ray must leave the Voronoi cell and so for some t', the point $c' = r(t')$ is in $\text{Vor}(u) \cap \text{Vor}(v)$ for some u in M. The set $\text{Vor}(u) \cap \text{Vor}(v)$ is a $d-1$-dimension $\mathcal{V}$-polyhedron, and thus can be written as $\text{conv}(Q) + \text{cone}(H)$ where Q is a subset of D-ball centers and H is a set of normals of unbounded D-balls. So, by Carathéodory's Theorem, there are subsets $Q' \subset Q$ and $H' \subseteq H$ such that $c \in \text{conv}(Q') + \text{cone}(H')$ and $|Q'| + |H'| \leq d$. Without loss of generality, we may assume $\frac{v+u}{2}$ is the origin. For $q \in Q'$ or $h \in H'$ let B_q and B_h be the bounded and unbounded D-balls corresponding to q and h respectively. Fix any $x \in B$. Lemma 4.16.4 implies that $x^T(q - c) \geq 0$ for some $q \in Q'$ or $x^T h > 0$ for some $h \in H'$. In the former case, Lemma 4.16.6 implies that $x \in B_q$. In the latter case, $x \in B_h$, because 0 is on the boundary of B_h for all $h \in H$.

We now observe that $0 \in B_q$ for all $q \in Q'$, because $|q - v|^2 = |q|^2 + |v|^2 > |q - 0|^2$. So, the bounded D-balls all have a common intersection at 0 and in fact at a sufficiently small open ball U centered at 0. So the intersection of U with the relative interior of $\text{cone}(H')$ is contained in the common intersection of the chosen D-balls. □

4.17 Technical Lemmas for Size Bounds

The following useful lemma guarantees that the feature size of any mesh vertex is maximized at the time it is inserted. Such a fact would be trivial if not for the possibility to change the underlying domains.

Lemma 4.17.1. *If M is a hierarchical mesh constructed by the* NetMesh *algorithm, then for all $v \in M$,*

$$\max_i \max_{\substack{\Omega \in H_i: \\ v \in M_{i\Omega}}} \mathbf{f}_{P_i}^{\Omega}(v) = \mathbf{f}_{P_j}^{\Omega_j}(v),$$

where j is the insertion time of v and $\Omega_j \in H_j$ is the domain it is inserted into.

Proof. If $v \notin P$ then Ω_i is the unique domain such that $v \in M_{\Omega_i}$. However, $P_{j\Omega_j} \subseteq P_{i\Omega_i}$. Thus, $\mathbf{f}_{P_i}^{\Omega_i}(v) \leq \mathbf{f}_{P_j}^{\Omega_j}(v)$ for non-input points.

If $v \in P$, then it is possible for v to be in more than one M_Ω. Clearly, the input feature size of v will be maximized at the highest domain in the hierarchy that contains it, i.e. at the largest scale. At time j, this domain is Ω_j Just as with the Steiner point case, any changes to this highest level domain do not eliminate any input points and therefore $\mathbf{f}_{P_i}^{\Omega_i}(v) \leq \mathbf{f}_{P_j}^{\Omega_j}(v)$ in this case as well. □

Lemma 4.17.2. *Let M be a τ-quality hierarchical mesh constructed incrementally. Given two vertices $u, v \in M$, if u was inserted before v then $\lambda_v \leq \frac{|u-v|}{1-\varepsilon}$.*

Proof. Let i be the insertion time of v and let $\Omega \in H_i$ be the domain it was inserted into. So, $\lambda_v = \mathbf{f}_{M_i}^{\Omega}(v)$. If either $u \in M_{i\Omega}$ or $u \notin B_\Omega$ then $\lambda_v \leq |u - v|$. So we may assume $u \in B_\Omega \setminus M'_\Omega$ and thus for some $\Omega' \in \text{children}(\Omega)$, $u \in B_{\Omega'}$. Since v does not encroach on Ω', we have that $|c_{\Omega'} - v| \leq \frac{|u-v|}{1-\varepsilon}$. Moreover, $c_{\Omega'} \in M_\Omega$, so $\lambda_v \leq |c_{\Omega'} - v| \leq \frac{|u-v|}{1-\varepsilon}$ as desired. □

Lemma (Lemma 4.9.1)**.** *If M is a hierarchical mesh constructed incrementally that satisfies the insertion radius invariant, then M also satisfies the feature size invariant.*

Proof. Let Ω be any domain and let v be any vertex of M_Ω. Let u be the nearest neighbor of v in M_Ω. This implies that $\mathbf{f}_M^\Omega(v) = |u - v|$ and so it will suffice to prove $\mathbf{f}_P^\Omega(v) \leq K_v|u - v|$. Let i (respectively j) be the insertion time of v (respectively u) and let Ω_i (Ω_j respectively) be the domain it was inserted into. Let $K'_v \in \{K'_C, K'_S, K'_I\}$ be the appropriate constant depending on how v was inserted and similarly for K'_u. We choose K so that

$$K \geq \frac{\max\{K'_C, K'_S, K'_I\}}{1 - \varepsilon} + 1 \tag{4.28}$$

There are two cases to consider.

Case: u inserted before v:

$$\begin{aligned}
\mathbf{f}_P^\Omega(v) &\leq \mathbf{f}_{P_i}^{\Omega_i}(v) && \text{[by Lemma 4.17.1]} \\
&\leq K'_v \lambda_v && \text{[by assumption]} \\
&\leq K'_v \frac{|u - v|}{1 - \varepsilon} && \text{[by Lemma 4.17.2]} \\
&< K|u - v|. && \text{[by (4.28)]}
\end{aligned}$$

Case: v inserted before u:

$$\begin{aligned}
\mathbf{f}_P^\Omega(v) &\le \mathbf{f}_P^\Omega(u) + |u-v| && \left[\mathbf{f}_P^\Omega \text{ is 1-Lipschitz}\right] \\
&\le \mathbf{f}_{P_j}^{\Omega_j}(u) + |u-v| && [\text{by Lemma 4.17.1}] \\
&\le K'_u \lambda_u + |u-v| && [\text{by assumption}] \\
&\le K'_u \frac{|u-v|}{1-\varepsilon} + |u-v| && [\text{by Lemma 4.17.2}] \\
&\le K|u-v|. && [\text{by (4.28)}]
\end{aligned}$$

□

Lemma 4.17.3. *Let Ω be any domain in the output. There exists an ordering $p_1 \ldots, p_j$ of the vertices of P_Ω such that for each $i = 3 \ldots j$, $\mathbf{f}_{P_i}^\Omega(p_i)/\mathbf{f}_{P_{i-1}}^\Omega(p_i) \ge \frac{12}{\varepsilon^3} + 1$, where $P_i = \{p_1, \ldots, p_i\}$.*

Proof. The desired ordering can be found by starting with the two farthest points of P_Ω as p_1 and p_2 followed by greedily adding any point that satisfies the desired property. Suppose this process gets stuck after adding i points and some $j - i$ points are leftover. Let $p \in P_i$ be such that some leftover point lies in $\mathrm{Vor}_{P_i}(p)$ and let q be the farthest such point from p. Let p' be the nearest neighbor of p in P_i. Let $r = \mathbf{f}_{P_i^\Omega}(q) = |p - q|$. Since the ordering was stuck, it must be that $|q-p'|/|q-p| \ge \frac{12}{\varepsilon^3}+$. By the triangle inequality, $|p'-p| \ge |p'-q| - |p-q|$. Combining the previous two statements give that $|p'-p| \ge \frac{12r}{\varepsilon^3}$. By our choice of p' as the nearest neighbor of p, we get that $\mathbf{annulus}(p, 2r, \frac{6r}{\varepsilon^3})$ is contained in $\mathrm{Vor}_{P_i}(p)$. Moreover, by our choice of q, this annulus is empty of points from P_Ω, contradicting Lemma 4.9.3. □

Lemma 4.17.4. *Let q and q' be any two input points and let r be the distance between them. If $A = \mathbf{annulus}(q, 2r, \frac{6r}{\varepsilon^3})$ contains no input points, then q and q' are inside some cage contained in A for all intermediate meshes after each has been inserted.*

Proof. Let $p_1, \ldots, p_k$ be all input points in $\mathbf{ball}(p, 2r)$ ordered by the order in which they were inserted. Clearly q and q' are among the p_i's. We will proceed by induction on k. In the base case, there are only two points, p_1 and p_2. Let Ω be the domain into which p_2 was inserted and P' and M' be the input and mesh vertices in the domain just after insertion. Since A contains no input points and $\mathbf{f}_P^\Omega$ is Lipschitz, $\mathbf{f}_P^\Omega(p_1) \ge \frac{6r}{\varepsilon^3} - 2r \ge \frac{4r}{\varepsilon^3}$. So, by Lemma 4.9.2, $\mathbf{f}_M^\Omega(p_1) \ge \frac{4r}{K\varepsilon^3}$. Since the algorithm chooses $\varepsilon < \frac{1}{K}$, we have that $\mathbf{f}_M^\Omega(p_1) \ge \frac{4r}{\varepsilon^2}$ and therefore, using the Lipschitz property, $\mathbf{f}_M^\Omega(p_2) \ge \mathbf{f}_M^\Omega(p_1) - |p_1 - p_2| \ge 4r(\frac{1}{\varepsilon^2} - 1)$. After insertion, we have that $\mathbf{f}_{M'}^\Omega(p_2) \le 4r$, because $p_1 \in M'_\Omega$. So, the ratio of nearest to second nearest neighbor distances for p_2 is bounded as

$$\frac{\mathbf{f}_{M'}^\Omega(p_2)}{\mathbf{f}_M^\Omega(p_2)} \le \frac{4r}{4r(\frac{1}{\varepsilon^2} - 1)} < \varepsilon.$$

Thus, the algorithm adds a cage around p_1 and p_2 as desired.

The inductive step has two cases. Either the ith point is added inside or outside the cage surrounding the first $i-1$ points (guaranteed to exist by induction). The latter case is identical to the base case, so it only remains to consider the case where p_i lies inside the cage from the previous round. If p_i does not encroach this cage, then it remains and we are done. If p_i does encroach this cage, then it will grow by a factor of $\frac{1}{\varepsilon}$. However, it's total size cannot exceed $\frac{1}{\varepsilon^2}$ times the distance from p_i to the center because otherwise it would not encroach. This distance is at most $4r$, so the cage is in A. Thus, the grown cage also satisfies the induction hypothesis and we are done. □

4.18 Technical lemmas regarding the feature size function

The following lemma extends the feature size invariant to mesh vertices in $B_\Omega \setminus M_\Omega$.

Lemma 4.18.1. *If (M, H) is a hierarchical mesh of an input set P such that the feature size invariant holds and no domain is ε-encroached then for all domains $\Omega \in H$ then for all $u \in M \cap B_\Omega$,*

$$\mathbf{f}_P^\Omega(u) \leq \alpha \mathbf{f}_M^\Omega(u),$$

where α is a constant depending only on the meshing parameters.

Proof. Fix a domain Ω and a vertex $u \in M \cap B_\Omega$. If $u \in M_\Omega$, then the feature size invariant implies the desired result. So, we may assume that $u \notin M_\Omega$. Let v be the nearest neighbor of u in M_Ω. Since u does not ε-encroach on any domains, we have that

$$|u - v| \leq \varepsilon \mathbf{f}_M^\Omega(v). \tag{4.29}$$

So,

$$\mathbf{f}_M^\Omega(v) \leq \frac{1}{1-\varepsilon} f_M^\Omega(u), \tag{4.30}$$

because $\mathbf{f}_M^\Omega$ is Lipschitz.

$$\begin{aligned}
\mathbf{f}_P^\Omega(u) &\leq \mathbf{f}_P^\Omega(v) + |u - v| && \left[\mathbf{f}_P^\Omega \text{ is 1-Lipschitz}\right] \\
&\leq K\mathbf{f}_M^\Omega(v) + |u - v| && \text{[by the feature size invariant]} \\
&\leq (K + \varepsilon)\mathbf{f}_M^\Omega(v) && \text{[by (4.29)]} \\
&\leq \frac{K + \varepsilon}{1 - \varepsilon}\mathbf{f}_M^\Omega(u). && \text{[by (4.30)]}
\end{aligned}$$

□

For quality meshes, it is possible to extend the feature size invariant to all points in the plane.

Lemma 4.18.2. *If M is a τ'-quality mesh of an input set P such that the feature size invariant holds and no domain is ε-encroached then for all domains $\Omega \in H_M$ then for all $x \in B_\Omega$,*

$$\mathbf{f}_P^\Omega(u) \leq \beta \mathbf{f}_M^\Omega(u),$$

where β is a constant depending only on the meshing parameters.

Proof. Fix a domain Ω and a point $x \in B_\Omega$. Let u be the nearest neighbor of x in M. Observe that $u \in B_\Omega$. So by the Lipschitz property of $\mathbf{f}_P^\Omega$ and Lemma 4.18.1,

$$\mathbf{f}_P^\Omega(x) = \alpha \mathbf{f}_M^\Omega(u) + |u - x|.$$

Because $\mathbf{f}_M^\Omega$ is also 1-Lipschitz,

$$\mathbf{f}_P^\Omega(x) = \alpha \mathbf{f}_M^\Omega(x) + (1 + \alpha)|u - x|.$$

Because we chose u to be the nearest neighbor, the diametral ball of u and x is empty and centered in B_Ω. Thus, Lemma 4.11.3 implies that

$$|u - x| \leq 24\tau'^2 \mathbf{f}_M^\Omega(x).$$

So, choosing $\beta = \alpha + (1 + \alpha)24\tau'^2$ we have that $\mathbf{f}_P^\Omega(x) \leq \beta \mathbf{f}_M^\Omega(x)$ as desired. □

4.19 Technical Lemmas for the Point Location Analysis

Lemma 4.19.1. *For all $i \in I$ and $x \in U_i$, $f_0(x) \leq c_{4.19.1} f_i(x)$, where $c_{4.19.1} = \frac{1}{1-e}$.*

Proof. Fix an index $i \in I$ and a point $x \in U_i$. Suppose for contradiction that $f_i(x) < (1 - e)f_0(x)$. We will show that $\mathbf{d}_{\mathcal{B}_M}(x, p_i) \leq 5$, contradicting the hypothesis that $i \in I$.

Recall that for $x \in U_i$, $f_i(x) \neq f_{i-1}(x)$. Combined with our supposition that $f_i(x) < (1-e)f_0(x)$, this implies that there can be at most one point of M_Ω in $\mathbf{ball}(x, \frac{1}{1-e}|x - p_i|)$. Call this point z and let Z denote the set of points of M whose ancestor in M_Ω is z.

We will construct a chain of balls $B_1, \ldots, B_6$ from p_i to x. To do this, the following claim is useful.

If $y \in \mathbf{ball}(x, |x - p_i|)$ does not encroach any domain of $H_{M'}$ then $\mathbf{ball}\left(\frac{x+y}{2}, \frac{|x-y|}{2}\right) \cap M \subseteq Z$.

We will give the construction using this claim and then give its proof.

Let B_1 be the maximal ball of $\mathcal{B}_M$ tangent to p_i centered on $\overline{xp_i}$. If $x \in B_1$, then $\mathbf{d}_{\mathcal{B}_M}(x, p_1) \leq 1$, so we may assume that $x \notin B_1$ and therefore $B_1 \subset \mathbf{ball}(x, |x - p_i|)$. So the claim implies that some vertex of Z is on the boundary of B_1.

Let B_4 be the maximal ball of $\mathcal{B}_M$ tangent to x centered on $\overline{xz}$. As with B_1, we may assume that $p_i \notin B_4$ and thus the claim implies some vertex of Z is on the boundary of B_4.

If $Z = \{z\}$ then $z \in \partial(B_1) \cap \partial(B_4)$ and thus $\mathbf{d}_{\mathcal{B}_M}(x, p_i) \leq 1$. So, we may assume that there is some domain $\Omega' \in H_{M'}$ centered at z whose parent is Ω. Let r' be the radius of Ω'. Since Ω' is not encroached and $\varepsilon < \frac{1}{3}$, $\mathbf{annulus}(z, r', 3r')$ is empty of points of M. Any two points a, b of such an annulus have $\mathbf{d}_{\mathcal{B}_M}(a, b) \leq 3$. This is easily seen by considering the plane through a, b, and z and

packing the annulus with balls of radius r' as in Figure 4.4. Since both B_1 and B_6 intersect the annulus, we may choose $a \in B_1$ and $b \in B_6$ and let $B_2, \ldots, B_5$ be the balls packing the annulus on the shortest path from a to b. The balls $B_1, \ldots, B_6$ witness that $\mathbf{d}_{\mathcal{B}_M}(x, p_i) \leq 5$ as desired.

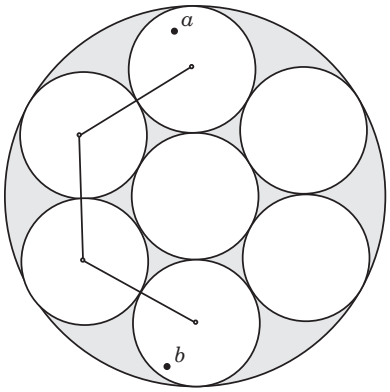

Figure 4.4: Any two points $a, b \in \mathbf{annulus}(z, r', 3r')$ have $\mathbf{d}_{\mathcal{B}_M}(a, b) \leq 3$.

To conclude the proof, we need only prove the claim. Let $B' = \mathbf{ball}(c, r)$ be the ball under consideration where $c = \frac{x+y}{2}$ and $r = \frac{|x-y|}{2}$. Suppose for contradiction that there exists $w \in B' \cap (M \setminus Z)$. Then, $|c - w| < r = |x - c|$. Let v be the ancestor of w in M_Ω and note that $v \neq z$. Observe that $|x - y| \leq (1 - \varepsilon)|x - v|$ for otherwise $f_0(x) \leq |x - v| \leq \frac{1}{1-\varepsilon} f_i(x)$ contrary to our supposition. It follows that

$$|x - v| \geq \frac{2\varepsilon r}{1 - \varepsilon}. \tag{4.31}$$

We can now bound $|v - w|$ in terms of r as follows.

$$\begin{aligned} |v - w| &\leq \varepsilon |y - v| && [y \text{ does not encroach}] \\ &\leq \varepsilon(|y - c| + |c - w| + |w - v|) && [\text{by the triangle inequality}] \\ &< \frac{2\varepsilon r}{1 - \varepsilon}. && [|y - c| = r \text{ and } |c - w| < r] \end{aligned} \tag{4.32}$$

This allows us to derive the following contradiction.

$$\begin{aligned} |c - w| &\geq |x - v| + |v - w| - |c - x| && [\text{by the triangle inequality}] \\ &> \left(\frac{2r}{1 - \varepsilon} - \frac{2\varepsilon}{1 - \varepsilon} - 1 \right) r && [\text{by (4.31) and (4.32)}] \\ &= r. \end{aligned}$$

□

Lemma 4.19.2. $\sum_{i \in I} (h_i - h_{i-1}) \leq (c_{4.19.2} - 1) h_0$.

Proof. First, we bound h_{jj} using Lemma 4.19.1 as follows.

$$h_{jj} = \int_{U_j} \frac{dx}{f_j(x)^d} \leq c_{4.19.1}^d \int_{U_j} \frac{dx}{f_0(x)^d} = c_{4.19.1}^d h_0. \tag{4.33}$$

For any $i \in I$, define i^* to be the largest element of I_0 less than i. If $i, j \in I$ and $i > j$, then $h_{ij} - h_{(i-1)j} = 0$ as guaranteed by the definition of U_j. Because $f_i \leq f_{i'}$ for all $i \leq i'$, $h_{(i-1)j} \geq h_{i^*j}$ for all i. The desired bound is now proven as follows.

$$\begin{aligned}
\sum_{i \in I} (h_i - h_{i-1}) &\leq \sum_{j \in I_0} \sum_{i \in I : i \leq j} (h_{ij} - h_{(i-1)j}) \\
&\leq \sum_{j \in I_0} \sum_{i \in I : i \leq j} (h_{ij} - h_{i^*j}) && [i^* \leq i - 1] \\
&= \sum_{j \in I_0} (h_{jj} - h_{0j}) \\
&\leq \sum_{j \in I_0} (c_{4.19.1}^d - 1) h_{0j} && [\text{by } (4.33)] \\
&= (c_{4.19.1}^d - 1) h_0.
\end{aligned}$$

Choosing $c_{4.19.2} = c_{4.19.1}^d$ completes the proof. □

Lemma 4.19.3. $h_0 \leq c_{4.19.3}$.

Proof. There are four types of domains to consider: the smallest domain Ω such that $B \subset B_\Omega$, domains Ω such that $|M_\Omega| = 0$, domains Ω such that $|M_\Omega| = 1$, and domains Ω such that M_Ω contains an entire cage $C_{\Omega'}$ of some domain $\Omega' \in H_M$. In the first case, the result follows easily from Lemma 4.11.3. In the second case, $\mathbf{f}_M^\Omega = \infty$, and thus, the integral evaluates to 0. In the third case, it is easy to evaluate the integral directly using polar coordinates to find that it is constant.

The last case is the interesting one. We use the coarse bounds that $\mathbf{f}_M^\Omega(x) \geq \delta r_{\Omega'}$ for $(1 - \delta - \gamma) r_{\Omega'} \leq |x - c_\Omega| \leq 2 r_{\Omega'}$ and $\mathbf{f}_M^\Omega(x) \geq \frac{1}{2}|x - c_\Omega|$ for $|x - c_\Omega| > 2 r_{\Omega'}$. Integrating with polar coordinates centered at c_Ω yields an answer $O(\log \frac{r_\Omega}{r_{\Omega'}})$. Only a constant number of points in a round may cause Ω' to grow because all but one must lie in the Voronoi cell of C_Ω and thus they are all within a constant D-ball distance of one another. So, $\frac{r_\Omega}{r_{\Omega'}} = O(1)$ and thus the integral also evaluates to $O(1)$. □

Chapter 5

On Fat Voronoi Diagrams

A Voronoi diagram is β-fat if every cell C is contained in a ball at most β times larger than a ball that it contains. We explore the complexity of cells in fat Voronoi diagrams in $\mathbb{R}^d$ with care to understand the dependence on d. Using standard methods, we prove that the average number of neighbors of any cell is $(4\beta)^d$. We then generalize this approach to give a worst case bound of $(4\beta)^d \log \tau$ neighbors, where τ is the aspect ratio of the Voronoi cell as defined in the Voronoi Refinement meshing literature. Note that all of these bounds are independent of the number of input points. We prove a stronger result in the plane, namely, that no cell in a fat planar Voronoi diagram has more than $\frac{6\pi}{\arcsin \frac{1}{4\beta}}$ neighbors. This is in marked contrast to the situation for general planar Voronoi diagrams or even general fat complexes, both of which may have cells of linear complexity. These general upper bounds are complemented by a lower bound on the complexity of fat Voronoi diagrams in $\mathbb{R}^d$, which says that the total number of faces in the diagram is $2^{\Omega(d)} \left(\frac{\sqrt{d}}{\beta}\right)^d$. Interestingly, this shows that although the worst-case complexity of Voronoi diagrams is avoided by fat Voronoi diagrams, so too is the best case.

5.1 Introduction

Many geometric search data structures employ decompositions of space into fat cells to achieve faster search, better approximation for proximity queries, and lower total complexity (as measured by the number of faces) when compared to their non-fat alternatives. In this chapter, we explore the special case of fat Voronoi diagrams with a particular emphasis on bounding the complexity of the cells locally and understanding the dependence of the dimension, providing both upper and lower bounds.

We prove that no cell in a fat planar Voronoi diagram has more than $O(1)$ neighbors. We also prove a bound in higher dimensions on the local complexity that depends on a geometric property of the cells. These general upper bounds are complemented by a nontrivial lower bound on the complexity of fat Voronoi diagrams.

Let M be a set of n points in $\mathbb{R}^d$. Recall that the Voronoi cell of a point $v \in M$ is the set of

all points of $\mathbb{R}^d$ that have v as a nearest neighbor in M. The **in-ball** of $\mathrm{Vor}(v)$ is the largest ball contained in $\mathrm{Vor}(v)$. It is denoted b_v and its radius is r_v. The **out-ball** of $\mathrm{Vor}(v)$ is its smallest enclosing ball. It is denoted B_v and its radius is R_v. We say that $\mathrm{Vor}(v)$ is β**-fat** if $R_v/r_v \leq \beta$. We say that a Voronoi diagram is β-fat if every cell is β-fat.

Recall that the definition of aspect ratio is similar to fatness with the added constraint that the in-ball and out-ball are centered at v. Although bounded aspect ratio implies fatness, the converse does not hold (see Figure 5.1). It is known that for all fat complexes, the number of neighbors of a cell is $O(1)$ on average. However, it is possible to improve these bounds to worst-case guarantees when the complex is a Voronoi diagram. In Section 5.3, we prove that the total complexity of a β-fat Voronoi diagram is $O(n)$ and then generalize this to a bound on the complexity of the cells in any dimension, where the bound depends on the log of the aspect ratio of the cells. This generalizes the bounds for the complexity of bounded aspect ratio Voronoi diagrams used in mesh generation [MTTW99, HMP06].

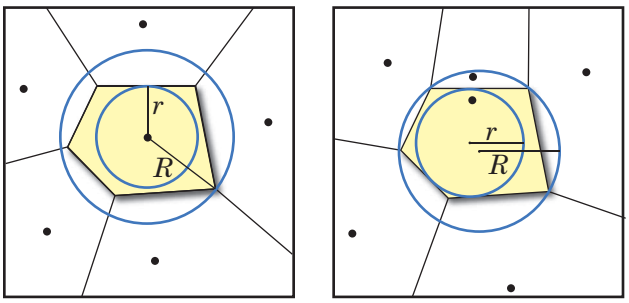

Figure 5.1: The cell on the left has good aspect ratio. The cell on the right is fat but it does not have good aspect ratio.

For the case of planar Voronoi diagrams, we prove a stronger bound, showing that no cell has more than a constant number of neighbors (see Section 5.4). Such a result does not hold for general planar Voronoi diagrams or for general fat complexes. There even exist simple examples where weighted fat Voronoi diagrams have cells with $O(n)$ neighbors (see Figure 5.2), so the result is not at all obvious. The proof method here is novel in that it exploits a tradeoff between angles in the Voronoi cell and angles in the Delaunay link (the polygon formed by the union of Delaunay triangles sharing a vertex). We conjecture that a similar bound holds for higher dimensions as well.

In Section 5.5, we turn our attention to lower bounds. We show that when β is a constant, the complexity of a cell in a fat Voronoi diagram is at least $2^{\Omega(d \log d)}$. Although an integer lattice gives rise to Voronoi cells with only $2^{O(d)}$ faces, the fatness of the cubical cells is only $\sqrt{d}$. This lower bound shows exactly the tradeoff between the fatness and the best-case complexity. Thus, fat Voronoi diagrams are bounded away from both the best-case and the worst-case complexity of Voronoi diagrams in general.

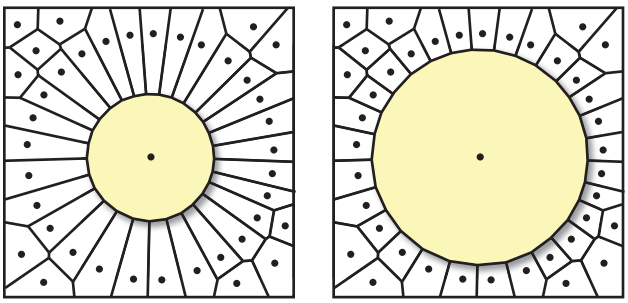

Figure 5.2: **Right:** A fat Voronoi cell can have arbitrarily many neighbors but at least some of them must be "skinny". **Left:** If weights are allowed on the vertices, the corresponding weighted Voronoi diagram can be fat and yet have unbounded maximum degree.

5.2 Related Work

The notion of *fat* objects is quite common in computational geometry. Though several different definitions exist in the literature, the differences are small (mostly just constant factors) for convex polytopes as in the case of Voronoi cells. These definitions arise naturally in settings where the bounds depend on packing arguments. In many cases, fatness is a convenience that permits better analyses, such as in the case of range searching and point location [OvdS94, dB95]. Fatness has also been considered for robot motion planning and ray shooting [vdSHO93, vdSO94, dB05, AdG08].

The problem of explicitly computing fat partitions of space was considered by Damian-Iordache [DI04]. Fatness is also a virtue when the complexes are used for point location and proximity search problems [DGK99, MM99, Gra10].

Erickson gives a thorough analysis of the blowup in Voronoi diagram complexity in low dimensions [Eri01] building on the classical results showing that Voronoi diagrams can have complexity $\Theta(n^{\lceil d/2 \rceil})$ by Klee [Kle80] and Seidel [Sei87]. Researchers have also looked at the complexity of Voronoi diagrams of random points, showing that this worst case complexity is not the norm. For example, Dwyer proved that n independent and identically distributed points sampled uniformly from a ball have a Voronoi diagram with $O(n)$ complexity [Dwy91]. This result was later extended by Bienkowski et al. to random samples from a cube [BDadHS05]. The linear complexity of Voronoi diagrams with bounded aspect ratio (a subset of fat Voronoi diagrams) is exploited in mesh generation [MTTW99] and is a primary motivation for the current research.

5.3 Fat Linear Complexity

The total complexity of a complex of fat objects has been studied in the context of robot motion planning [vdSHO93]. The trick to bounding the complexity is to bound the number of larger neighbors of any cell by volume packing arguments. As shown in the recent paper by Gray, this trick is directly applicable to the case of fat Voronoi diagrams [Gra10]. Thus, the number of $(d-1)$-dimensional faces of the Voronoi diagram is $2^{O(d)}n$ and the total complexity is $2^{O(d^2)}n$ counting

faces of all dimensions.[1] Although the methods are quite standard and the proofs are not difficult, we present the full proofs of these results in Theorems 5.3.1 and 5.3.2 below for three reasons: one, previous proofs use a different definition of fatness that can hide the dependence on the dimension, which is critical to our analysis; two, we prove slightly tighter bounds on the total complexity with respect to the dependence on d; and three, the proofs introduce the machinery necessary to prove a local complexity bound in Theorem 5.3.3.

We begin by ordering the vertices of M by the radius of their in-balls, so $u \prec v$ when $r_u < r_v$, and ties are broken arbitrarily.

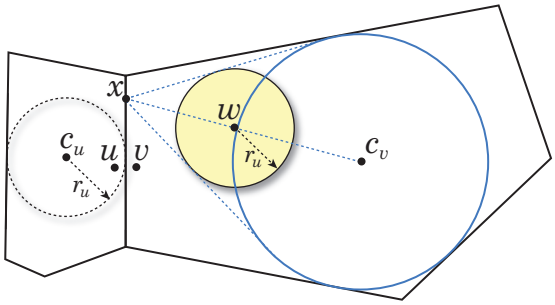

Figure 5.3: The ball at w has is contained in V_v and has radius r_u.

Theorem 5.3.1 (Fat Linear Complexity)**.** *Let M be a set of n points in $\mathbb{R}^d$. If Vor_M is β-fat, then the number of edges in Del_M is at most $(4\beta)^d|M|$.*

Proof. We will charge each edge to its smaller neighbor with respect to the $\prec$ relation. In this way, it will suffice to show that the number of larger neighbors of any vertex is $(4\beta)^d$. Let u be any vertex and let v be any Delaunay neighbor of u such that $u \prec v$ as in Figure 5.3. Let x be a point in $\mathrm{Vor}(u) \cap \mathrm{Vor}(v)$ and let c_u be the center of b_u, the in-ball of $\mathrm{Vor}(u)$. Since both x and c_u are in B_u, their distance is bounded as $|x - c_u| \leq (2\beta - 1)r_u$.

Let c_v denote the center of b_v. Since $\mathrm{Vor}(v)$ is convex and contains both v and b_v, for any $\alpha \in [0, 1]$, the ball centered at $w_\alpha = \alpha c_v + (1 - \alpha)x$ with radius αr_v is contained in $\mathrm{Vor}(v)$. In particular, if $\alpha = r_u/r_v$ and $w = w_\alpha$, then $\mathbf{ball}(w, r_u) \subset \mathrm{Vor}(v)$. Moreover, because x and c_v are contained in B_v,

$$|x - w| = \alpha|x - c_v| \leq \alpha 2R_v \leq 2\beta r_u.$$

So, by the triangle inequality, $|c_u - w| \leq (4\beta - 1)r_u$.

We complete the proof by the following volume packing argument. For any larger neighbor v of u, there is a corresponding w so that $\mathbf{ball}(w, r_u)$ is contained in $B = \mathbf{ball}(c_u, 4\beta r_u)$. Since these balls are disjoint (each is in a different Voronoi cell), their total volume cannot exceed the volume

[1]The actual bound proved by Gray is $2^{2^{O(d)}}$ [Gra10], however it is a simple exercise to tighten this bound by a slightly more careful analysis as shown in Theorem 5.3.2.

of B. So, the number of larger neighbors of u is at most

$$\frac{\mathbf{Vol}(B)}{\mathbf{Vol}(\mathbf{ball}(w_\alpha, r_u))} = \frac{(4\beta r_u)^d \mathbb{V}_d}{r_u^d \mathbb{V}_d} = (4\beta)^d.$$

□

This idea easily extends to give the bound on the number of faces of other dimensions.

Theorem 5.3.2. *Let M be a set of n points in $\mathbb{R}^d$. If Vor_M is β-fat, then the number of faces in Vor_M is at most $(4\beta)^{d^2} n$.*

Proof. Counting the faces of the Voronoi diagram is equivalent to counting simplices in the Delaunay triangulation. Using the total ordering $\prec$ on the vertices, we charge each simplex to its minimum vertex. So, each vertex u can only be charged for a simplex if that simplex contains only larger neighbors of u and u itself. By the preceding theorem, there are only $(4\beta)^d$ such neighbors so the total number of j-simplices charged to any vertex is $\binom{(4\beta)^d}{j}$. Summing over all values of j from 1 to d yields the desired bound. □

The preceding theorems imply an average-case bound on the complexity of the individual cells. In the rest of this chapter, we will explore the different ways these average-case bounds can be improved to worst-case bounds.

5.3.1 A Bound on the Local Complexity

In this section we adapt the proof technique from Section 5.3, to prove a worst-case bound on the number of neighbors of a cell in terms of both its aspect ratio and its fatness. We conjecture that the dependence on the aspect ratio is not necessary, but such a theorem has proved elusive. Still, the main theorem of this section generalizes the known results for Voronoi diagrams of bounded aspect ratio [MTTW99].

Theorem 5.3.3. *Let M be a set of n points in $\mathbb{R}^d$. If Vor_M is β-fat, then for any vertex $p \in M$, p has at most $(8\beta)^d \log \tau$ neighbors, where τ is the aspect ratio of $\mathrm{Vor}(p)$.*

Proof. Let $q_1, \ldots, q_t$ be the neighbors of p. For each q_i, let $\mathbf{ball}(c_i, r_i)$ be the in-ball of $\mathrm{Vor}(q_i)$ and let R_i be the radius of the out-ball of $\mathrm{Vor}(q_i)$. Let x_i be the point of $\mathrm{Vor}(p) \cap \mathrm{Vor}(q_i)$ closest to p. Let $\alpha_i = \frac{|x_i - p|}{2\beta r_i}$. Note that $\alpha_i < 1$ because

$$\alpha_i = \frac{|x_i - p|}{2\beta r_i} = \frac{|x_i - q|}{2\beta r_i} < \frac{2R_i}{2\beta r_i} \leq 1.$$

We define a new ball centered at a point $w_i = \alpha_i c_i + (1 - \alpha_i) x_i$ with radius $r_i' = \alpha_i r_i$. As before, the convexity of $\mathrm{Vor}(q_i)$ implies that $\mathbf{ball}(w_i, r_i') \subset \mathrm{Vor}(q_i)$.

We will show that w_i is not too far from p, but first we must show that w_i is not too far from x_i.

$$\begin{aligned}
|w_i - x_i| &= \alpha_i |x_i - c_i| && \text{[by the definition of } w_i\text{]} \\
&\leq \alpha_i (2R_i - r_i) && \text{[because the out-ball of Vor}(q) \text{ contains } \mathbf{ball}(c_i, r_i)\text{]} \\
&\leq \alpha_i (2\beta r_i) - \alpha_i r_i && \text{[because Vor}(q) \text{ is } \beta\text{-fat]} \\
&= |x_i - p| - r'_i. && \big[\text{by the definition of } \alpha_i \text{ and } r'_i\big]
\end{aligned}$$

Applying the triangle inequality, we get that $|p - w_i| \leq 2|p - x_i| - r'_i$. Thus, for each q_i, the ball $B_i = \mathbf{ball}(w_i, r'_i)$ is contained in $\mathbf{ball}(p, 2|x_i - p|)$. Any two such balls are disjoint because each is contained in a different Voronoi cell.

We partition the neighbors of p based on their distance to p into layers $L_j = \{q_i : s2^j \leq |x_i - p| < s2^{j+1}\}$, where $s = \min_{q_i} |p - x_i|$. We have shown that the balls $\{B_i : q_i \in L_j\}$ are all contained in $\mathbf{ball}(p, \max_{q_i \in L_j} 2|x_i - p|)$ which is contained in $\mathbf{ball}(p, s2^{j+2}))$. So, we may bound the volume of the larger ball by the sum of the volumes of the contained balls, since all are disjoint:

$$\mathbf{Vol}(\mathbf{ball}(p, s2^{j+2})) > \sum_{q_i \in L_j} \mathbf{Vol}(\mathbf{ball}(w_i, r'_i)).$$

Therefore,

$$(s2^{j+2})^d \geq \sum_{q_i \in L_j} (r'_i)^d \geq \left(\frac{s2^j}{2\beta}\right)^d |L_j|.$$

It follows that $|L_j| \leq (8\beta)^d$. There are at most $\log \tau$ such levels L_j, so there are at most $(8\beta)^d \log \tau$ neighbors of p. □

Note that this proof requires that the complex is a Voronoi diagram. Indeed, the theorem is false without this assumption, unlike the preceding theorems which apply (in slightly modified form) for general fat complexes.

5.4 Local Complexity in the plane

In this section, we prove that the cells of planar fat Voronoi diagrams have constant complexity. Let Vor_M be a fat Voronoi diagram. For any $p \in M$, let Q_p be the set of Delaunay neighbors of p. That is, for each $q \in Q_p$, there is a Voronoi edge (u, v) dual to the Delaunay edge (p, q). For input points in general position, there is a unique edge e_u (e_v) in the Voronoi diagram that emanates from u (respectively v) but is not a boundary edge of Vor(p). Let ℓ_u and ℓ_v be the lines containing e_u and e_v respectively. We define two relevant angles:

1. $\theta_q :=$ the angle between ℓ_u and ℓ_v.

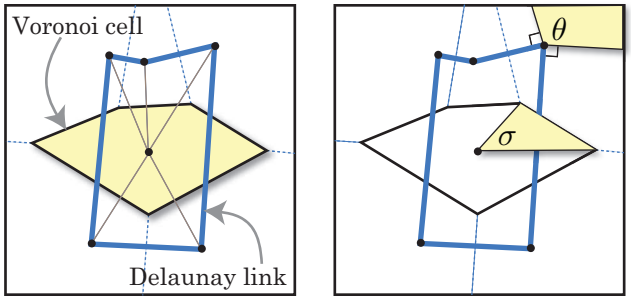

Figure 5.4: The proof of Theorem 5.4.5 uses the Voronoi cell and the dual Delaunay link. There is a natural tradeoff between the angle σ subtended from p by an edge of the Voronoi cell and θ, the angle between the normals of the dual Delaunay edges. Although, the Voronoi cell is convex, the Delaunay link may not be so one must take care to maintain signs of angles.

2. $\sigma_q := \angle vpu = \angle uqv$.

We observe the following fact about these angles.

$$\sum_{q \in Q_p} \theta_q \leq 2\pi \text{ and } \sum_{q \in Q_p} \sigma_q \leq 2\pi. \tag{5.1}$$

The first inequality sums the change in the normals along the Delaunay link of p which is a simple closed curve and thus yields 2π, possibly omitting the non-negative angle at p. The second inequality follows directly from the convexity of $\text{Vor}(p)$. In both cases equality holds for Voronoi cells that do not intersect the boundary of the bounding region. This gives us two ways to measure the angles, one from the perspective of the Delaunay triangulation and the other from the perspective of the Voronoi diagram. Ultimately, we will show that at least one of σ_q or θ_q must be larger than a fixed constant for each neighbor q. This will allow us the bound the size of Q_p.

Let $q \in Q_p$ be any neighbor of p and let (u, v) be the Voronoi edge dual to (p, q) as illustrated in Figure 5.5. Assume without loss of generality that $|q - u| \leq |q - v|$. To simplify notation, let $\sigma = \sigma_q$ and let $\theta = \theta_q$. Let r be the radius of the in-ball of $\text{Vor}(q)$. Let h be the distance from v to ℓ_u. We bound h in the following lemma.

Lemma 5.4.1. $h \leq 2\beta r \sin\sigma$

Proof. Let q' be the point on ℓ_u such that $\angle uq'v = \sigma$. We see that $h = |q' - v| \sin\sigma$. Thus, it will suffice to bound $|q' - v|$. Let S be the circumcircle of u, v, and q and let c be its center. The point q' is also on this circle because $\angle uqv = \angle uq'v = \sigma$.

If c is inside the triangle $\triangle uqv$ then we observe that this triangle is acute and its smallest enclosing ball is S. The radius of the smallest enclosing ball of $\text{Vor}(q)$ must be at least the radius of S because $\triangle uqv \subset \text{Vor}(q)$. Therefore, $|q - v'| \leq 2|q - c| \leq 2\beta r$.

If c is outside the triangle $\triangle uqv$ then the asumption that $|q-u| \leq |q-v|$ implies that $|q'-v| \leq |q - v| \leq 2\beta r$. So, we conclude that $|q' - v| \leq 2\beta r$ and thus, $h \leq 2\beta r \sin\sigma$ as desired. □

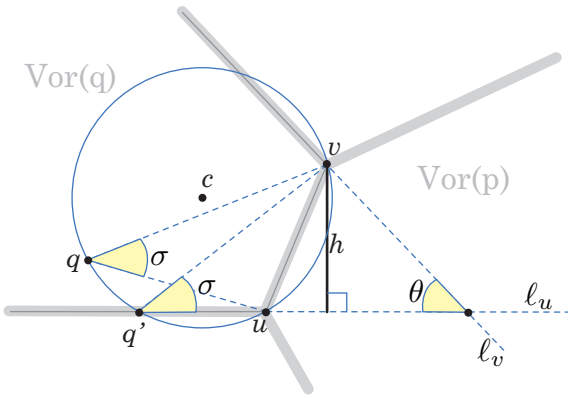

Figure 5.5: An illustration of the proofs of the lemmas on the local complexity of fat Voronoi diagrams in the plane. The trick is to compare σ and θ by comparing each to h.

Lemma 5.4.2. *If* $\theta \geq 0$ *then* $\max\{\sigma, \theta\} \geq \arcsin \frac{1}{4\beta}$.

Proof. Since $\beta > 1$ and $\sigma > 0$, the result is trivial for $\theta \geq \frac{\pi}{2}$. So, we may assume that $\cos\theta > 0$. Let c be the center of the in-ball of $\mathrm{Vor}(q)$. Clearly, c has distance at least r from ℓ_u. Let z be the point on this ball closest to ℓ_v. So the distance from z to ℓ_u is at least $r + r\cos\theta$, and since $\cos\theta > 0$,

$$\mathbf{d}(z, \ell_u) \geq r. \tag{5.2}$$

Since z is in the in-ball, it is also in $\mathrm{Vor}(q)$ and thus

$$|z - v| \leq 2\beta r. \tag{5.3}$$

So, we derive the following lower bound on h.

$$\begin{aligned} h &= \mathbf{d}(v, \ell_u) && \text{[by definition]} \\ &\geq \mathbf{d}(z, \ell_u) - |z - v| \sin\theta && \text{[by the triangle inequality]} \\ &\geq r - 2\beta r \sin\theta. && \text{[by the inequalities (5.2) and (5.3)]} \end{aligned}$$

Applying the upper bound on h from Lemma 5.4.1, we get that

$$r - 2\beta r \sin\theta \leq h \leq 2\beta r \sin\sigma,$$

and thus,

$$\frac{1}{2\beta} \leq \sin\sigma + \sin\theta.$$

The above inequality implies the lemma because $0 \leq \theta, \sigma \leq \pi$. □

Lemma 5.4.3. *If* $\theta < 0$ *then* $\sigma \geq \arcsin \frac{1}{\beta}$.

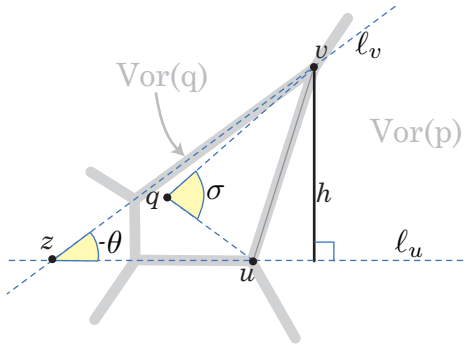

Figure 5.6: When $\theta < 0$, the intersection of ℓ_u and ℓ_v lies on the other side of the line segment $\overline{uv}$. Since these lines are support lines for the Voronoi cell of q, it is clear that $\sigma \geq -\theta$.

Proof. When $\theta \leq 0$ we have $h \geq 2r$, which, combined with Lemma 5.4.1, implies

$$2r \leq 2\beta r \sin \sigma.$$

The statement follows directly from this inequality. □

Lemma 5.4.4. *If $\theta < 0$ then $\sigma + \theta \geq 0$.*

Proof. Let $w = \ell_u \cap \ell_v$. Since $\theta < 0$, w is on the same side of the line segment $\overline{uv}$ as q. Moreover, all of $\mathrm{Vor}(q)$ and in particular q is contained in $\triangle uvw$ (see Figure 5.6). It follows that

$$\sigma = \angle uqv \geq \angle uwv = -\theta,$$

and so we get the desired result by adding θ to both sides. □

We are now ready to prove the main theorem of this section. It states that the local complexity of fat Voronoi diagrams in the plane is a constant that depends only on the fatness of the Voronoi diagram.

Theorem 5.4.5. *If Vor_M is β-fat then for all $p \in M$, p has at most $\frac{6\pi}{\arcsin \frac{1}{4\beta}}$ Delaunay neighbors.*

Proof. Let p be any vertex and let Q be the set of Delaunay neighbors of p. Partition Q based on the magnitude of θ_q into

$$Q^- = \{q \in Q : \theta_q < 0\} \text{ and}$$
$$Q^+ = Q \setminus Q^-.$$

We will bound the size of each set individually. Using Lemma 5.4.3, we derive the following.

$$\begin{aligned} 2\pi &\geq \sum_{q \in Q} \sigma_q && \text{[by Equation (5.1)]} \\ &\geq \sum_{q \in Q^-} \sigma_q && \text{[because } \sigma_q \geq 0 \text{ for all } q \in Q\text{]} \\ &\geq |Q^-| \arcsin \frac{1}{\beta}. && \text{[by Lemma 5.4.3]} \end{aligned}$$

This implies that $|Q^-| \leq \frac{2\pi}{\arcsin \frac{1}{\beta}}$. We now bound $|Q^+|$ as follows.

$$\begin{aligned} 4\pi &\geq \sum_{q \in Q} \sigma_q + \theta_q && \text{[by Equation (5.1)]} \\ &\geq \sum_{q \in Q^+} \sigma_q + \theta_q && \text{[by Lemma 5.4.4]} \\ &\geq \sum_{q \in Q^+} \max\{\sigma_q, \theta_q\} && \text{[because } \sigma_q, \theta_q \geq 0\text{]} \\ &\geq |Q^+| \arcsin \frac{1}{4\beta}. && \text{[by Lemma 5.4.2]} \end{aligned}$$

This implies that $|Q^+| \leq \frac{4\pi}{\arcsin \frac{1}{4\beta}}$. So we see that

$$|Q| = |Q^-| + |Q^+| \leq \frac{2\pi}{\arcsin \frac{1}{\beta}} + \frac{4\pi}{\arcsin \frac{1}{4\beta}} < \frac{6\pi}{\arcsin \frac{1}{4\beta}}.$$

□

The preceding proof does not use the fact that the Voronoi cell of p is fat, only that all of its neighbors are fat. In fact, one could adapt this proof to show that any Voronoi diagram in which the neighbors of any skinny cell are all fat has constant local complexity. This same notion was encountered by Maneewongvatana and Mount in their work on kd-trees [MM99]. In that setting, they analyzed a splitting heuristic for kd-trees and showed that although it doesn't guarantee all cells are fat, it does achieve similar properties to fat decompositions because the neighbors of skinny cells are all fat.

5.5 Lower bounds

In this section, we show that any β-fat Voronoi Diagram has at least $2^{\Omega(d \log d)} n$ faces when β is a constant independent of d. The lower bound also shows how the complexity varies when the fatness depends on the dimension. A nice example of this occurs when the point set is an integer lattice. In this case, the local complexity is only 2^d, but this does not violate the lower bound because high

dimensional cubes are not fat; their fatness is only $\sqrt{d}$. The theorem shows that the fatness must reach $\Omega(\sqrt{d})$ in order to reduce the local complexity to $2^{O(d)}$.

Theorem 5.5.1. *If* Vor_M *is a* β*-fat Voronoi diagram in* $\mathbb{R}^d$*, then every vertex of* M *has at least* $2^{\Omega(d)}\left(\frac{\sqrt{d}}{\beta}\right)^d$ *faces in its Voronoi cell.*

We postpone the proof until we have established the necessary definitions and a lemma. We start with a couple standard definitions from convex geometry.[2] For any vertex $p \in M$, let $s_1 \dots s_k$ be the simplices of Del_M with a vertex at p. For any set of points A, recall that $\mathrm{cone}(A)$ denotes the nonnegative, linear combinations of points of A (summing coordinate-wise),

$$\mathrm{cone}(A) = \left\{ \sum_{a_i \in A} c_i a_i : c_i \geq 0 \right\}.$$

The **polar** A° of a set $A \subset \mathbb{R}^d$ is defined as.

$$A^\circ = \{y \in \mathbb{R}^d : y^T a \leq 1 \text{ for all } a \in A\}.$$

Let s be any simplex with a vertex at p and let $C_s = \mathrm{cone}(\{p - q : q \in s\})$. There is a circumcenter x dual to s at a vertex of $\mathrm{Vor}(p)$. Let $C_x = \{y - x : |y - p| \leq |y - q| \text{ for all } q \in s\}$. Note that $C_s = C_x^\circ$.

Let b be the in-ball of $\mathrm{Vor}(p)$. If we let B_x denote $\mathrm{cone}(\{y - x : y \in b\})$, then we have $B_x \subset C_x$. Polarity reverses containment, so $C_x^\circ \subset B_x^\circ$. Letting B_s be the cone polar to B_x, it follows that

$$C_s = C_x^\circ \subset B_x^\circ = B_s. \tag{5.4}$$

The proof of the Lower Bound Theorem requires the following lemma whose purpose is to bound the fraction of a ball's volume that a Delaunay simplex may cover in terms of the dimension and the fatness of the Voronoi diagram.

Lemma 5.5.2. *Let* Vor_M *be a* β*-fat Voronoi diagram in* $\mathbb{R}^d$ *and let* p *be a point of* M*. For all simplices* s *at* p*,* $\mathbf{Vol}(C_s \cap \mathcal{B}) < \frac{\mathbb{V}_d}{2^{\Theta(d)}\left(\frac{\sqrt{d}}{\beta}\right)^d}$*, where* $\mathcal{B}$ *is the unit ball centered at the origin.*

Proof. Let H_1 and H_2 be the halfspaces normal to the axis of B_s at distances $\frac{1}{2\beta}$ and 1 respectively that contain the origin. Since $\mathcal{B} \subset H_2$, it follows that $\mathbf{Vol}(C_s \cap \mathcal{B}) \leq \mathbf{Vol}(C_s \cap H_2)$. The two sets $C_s \cap H_2$ and $B_s \cap H_2$ are each the convex closure of a base and a vertex at the origin. Moreover, their bases both lie in the boundary of H_2 so the ratio of their volumes is equal to the ratio of the $d-1$-dimensional volumes of their bases. Since the base of $C_s \cap H_s$ is a $d-1$-simplex contained in the $d-1$-dimensional ball that is the base of $B_s \cap H_2$, the ratio of their volumes is bounded using

[2]The reader is encouraged to read Barvinok [Bar02] for a nice treatment of convexity theory.

Stirling's approximation as follows.

$$\frac{\mathbf{Vol}(C_s \cap H_2)}{\mathbf{Vol}(B_s \cap H_2)} \leq \frac{\mathbf{Vol}(\mathbb{S})}{\mathbb{V}_{d-1}} \leq \frac{1}{2^{\Theta(d)} d^{\frac{d}{2}}},$$

where $\mathbb{S}$ represents the regular $d-1$-simplex inscribed in the $d-1$-dimensional unit ball. By scaling $B_S \cap H_2$ down by a factor of 2β, we get $B_S \cap H_1$, and therefore

$$\mathbf{Vol}(B_S \cap H_2) = (2\beta)^d \mathbf{Vol}(B_S \cap H_1).$$

Since $B_S \cap H_1 \subset \mathcal{B}$, we have that $\mathbf{Vol}(B_S \cap H_1) < \mathbb{V}_d$. Combining the preceding statements, we get

$$\mathbf{Vol}(C_s \cap \mathcal{B}) < \mathbf{Vol}(C_s \cap H_2) \leq \frac{\mathbf{Vol}(B_s \cap H_2)}{2^{\Theta(d)} d^{\frac{d}{2}}} = \frac{\mathbf{Vol}(B_s \cap H_1)}{2^{\Theta(d)} \left(\frac{\sqrt{d}}{\beta}\right)^d} < \frac{\mathbb{V}_d}{2^{\Theta(d)} \left(\frac{\sqrt{d}}{\beta}\right)^d}.$$

□

5.5.1 Proof of Theorem 5.5.1

Proof. Let p be any vertex and let s be a simplex of Del_M with a vertex at p and a circumcenter at x. Let θ_x and θ_s be the half-angles of the circular cones B_x and B_s respectively. Because the cones are polar to each other, $\theta_x + \theta_s = \frac{\pi}{2}$. Because $\mathrm{Vor}(p)$ is β-fat, $\theta_x \geq \arcsin \frac{1}{2\beta}$. So,

$$\cos \theta_s \geq \cos\left(\frac{\pi}{2} - \arcsin \frac{1}{2\beta}\right) = \frac{1}{2\beta}. \tag{5.5}$$

Let $\mathcal{B}$ be the unit ball centered at the origin. The cones $\{C_{s_i}\}_{i=1}^k$ have disjoint interiors and cover $\mathbb{R}^d$. Applying Lemma 5.5.2 and a simple volume packing, we get that

$$\mathbb{V}_d = \sum_{i=1}^{k} \mathbf{Vol}(C_{s_i} \cap \mathcal{B}) \leq \frac{k\mathbb{V}_d}{2^{\Theta(d)} \left(\frac{\sqrt{d}}{\beta}\right)^d}. \tag{5.6}$$

It follows that k, the number of simplices at p, is at least $2^{\Omega(d)} \left(\frac{\sqrt{d}}{\beta}\right)^d$. □

5.6 Concluding Remarks

We have presented a theory of fat Voronoi diagrams with an emphasis on bounding their complexity. The original motivation for this work was for problems in mesh generation, and indeed this is an area of ongoing work. However, we contend that bounding the complexity of these fat Voronoi diagram is interesting on its own both as a special case of classic problems of bounding polytope complexity and for its relation to fat complexes used in geometric search problems. We have given a proof of the constant local complexity of fat Voronoi diagrams in the plane, a parameterized

bound in higher dimensions, and a nontrivial lower bound.

Chapter 6

Geometric Persistent Homology

6.1 Overview

In this chapter, we present a summary of basic definitions and results in the theory of persistent homology. We start with topological definitions in Section 6.2.2, presenting topological spaces, simplicial homology, homotopy equivalences, and singular homology. Then, in Section 6.3, we define persistent homology and describe a basic algorithm for computing it. Next, we connect the combinatorial theory to a geometric one in Section 6.4. We finish this introductory chapter with an explanation of stability and approximation of persistence in Section 6.5.

6.2 Topology

In this section, we give the basic topological definitions that will be needed later. It is included here for completeness, but the reader who is already familiar with topology may skip ahead.

6.2.1 Topological Spaces

A **topological space** is a set X together with a family $\mathcal{O}$ of subsets called **open** sets satisfying the following conditions:

1. $\emptyset \in \mathcal{O}$,
2. $X \in \mathcal{O}$,
3. for any $Y \subset \mathcal{O}$, $\bigcup_{S \in Y} S \in \mathcal{O}$, and
4. for any finite $Y \subset \mathcal{O}$, $\bigcap_{S \in Y} S \in \mathcal{O}$.

The most important topological space for this thesis is the **standard topology on** $\mathbb{R}^d$ generated by the set of open balls, $\{\mathbf{ball}(c, r) : c \in \mathbb{R}^d, r \in \mathbb{R}_{\geq 0}\}$. We will also need to consider the **subspace topology**, which is defined on a subset $X \subset \mathbb{R}^d$ to be the family of sets of the form $Y \cap X$ for which Y is open in $\mathbb{R}^d$.

Abstract simplicial complexes do not have any geometry, however they can be embedded in Euclidean space. An abstract simplicial complex with n vertices has a natural embedding in $\mathbb{R}^n$ as a subcomplex of the **standard simplex** $\triangle_{n-1} = \text{conv}\{e_1, \ldots, e_n\}$, where $e_1, \ldots, e_n$ are the standard basis vectors for $\mathbb{R}^n$. Thus, when we speak of the topology of an abstract simplicial complex, we are referring to the subspace topology it inherits from $\mathbb{R}^n$ in this embedding.

6.2.2 Simplicial Homology

There are several homology theories to choose from. We present simplicial homology over the field $\mathbb{Z}_2$ of integers mod 2, which is particularly relevant to the situation often arising in computational problems where topological spaces are represented by simplicial complexes. Assuming the coefficients are in a field guarantees that the homology groups defined below are all vector spaces. In general, homology may be defined over arbitrary rings. Our approach will be to first present simplicial homology in purely a combinatorial and algebraic way. Then, we will show how this combinatorial approach is a special case of a more general approach using in linear algebra. Finally, we will connect the theory of simplicial homology to a more general theory of singular homology, which we will occasionally need later in the thesis. Though our treatment here is far from complete, it will provide the necessary definitions and notation to understand the main results of Chapters 7 and 8. For a thorough treatment of algebraic topology, we suggest the book by Hatcher [Hat01]. For treatments directed more at a computer science audience, the books by Zomorodian [Zom09] or Edelbrunner and Harer [EH09] are recommended. Lastly, for a combinatorial introduction to topology, see the book by Henle [Hen79].

Let $\mathcal{K}$ be an abstract simplicial complex. A subset of k-simplices is a ***k*-chain**. A k-chain is a formal sum of simplices We form the ***k*th chain group** $\mathcal{C}_k$ by defining addition as the symmetric difference between sets. That is, for k-chains A and B, $A + B := (A \cup B) \setminus (B \cap A)$.

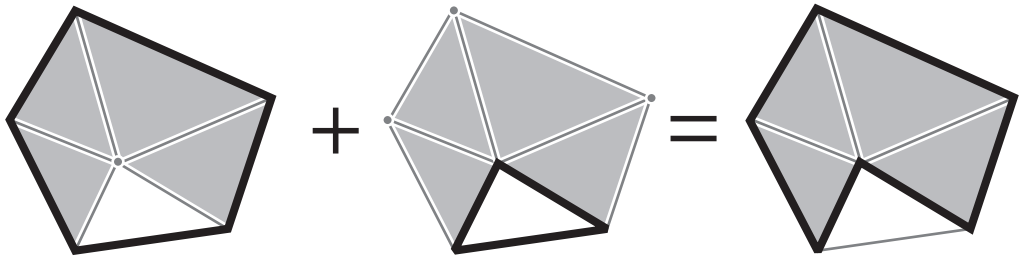

Figure 6.1: The addition of two 1-chains. The two cycles are homologous because their difference is the boundary of a 2-chain (i.e. a set of triangles).

A k-simplex $\sigma = \{v_0, \ldots, v_k\}$ has $k+1$ simplices of dimension $k-1$ on its boundary, denoted $\sigma_i = \sigma \setminus \{v_i\}$ for $i = 0 \ldots k$. The ***k*th boundary operator** ∂_k takes a k-simplex σ to the set of $(k-1)$-simplices on its boundary,

$$\partial_k \sigma = \sum_{i=0}^{k} \sigma_i$$

We extend ∂_k to all k-chains $c \in \mathcal{C}_k$ using addition,

$$\partial_k c = \sum_{\sigma \in c} \partial_k \sigma.$$

The boundary operator induces two natural subgroups of $\mathcal{C}_k$:

1. the ***k*th cycle group** $Z_k := \ker \partial_k = \{c \in \mathcal{C}_k : \partial_k c = 0\}$, and
2. the ***k*th boundary group** $B_k := \text{im}\partial_{k+1} = \{\partial c : c \in \mathcal{C}_{p+1}\}$.

Intuitively, the cycle group is the set of k-chains without boundary, like a cycle or set of cycles in a graph or 1-chain. The boundary group is the set of k-chains that are the boundaries of $(k+1)$-chains.

It is an easy exercise to show that $\partial_{k-1}\partial_k c = \emptyset$ for all k-chains c and therefore B_k is a subgroup of Z_k. Thus it is possible to define the ***k*th homology group** as

$$H_k = Z_k / B_k = \ker \partial_k / \text{im}\partial_{k+1}.$$

The rank of the kth homology group is ***k*th Betti number**.

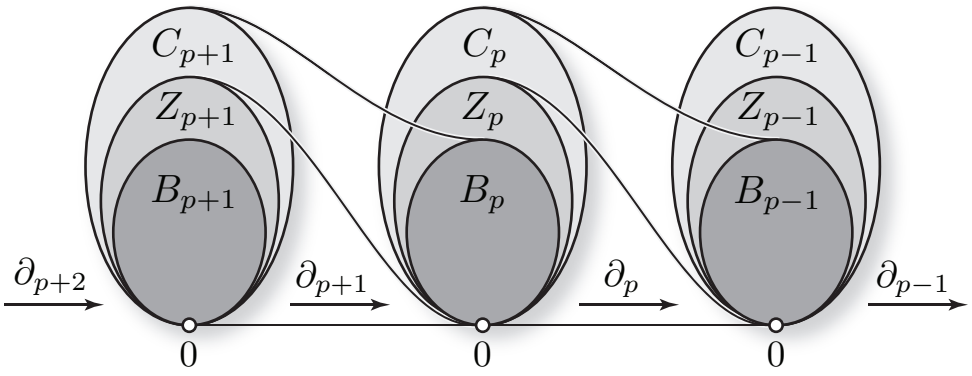

Figure 6.2: The relationship between the chain, cycle, and boundary groups is illustrated along with their boundary images.

Representing chains as characteristic vectors over $\mathbb{Z}_2$ gives a basis for the chain group vector space. Each boundary map ∂_k is a linear map and may be written as a matrix in $\mathbb{Z}_2^{n_{k-1} \times n_k}$, where n_i denotes the number of i-simplices in $\mathcal{K}$ for all i. The homology group H_k depends on both the underlying space ($\mathcal{K}$) and the coefficient ring ($\mathbb{Z}_2$), so it is sometimes written as $H_k(\mathcal{K}; \mathbb{Z}_2)$ or simply $H_k(\mathcal{K})$ when the coefficient ring is assumed.

6.2.3 Homotopy Equivalences

Let X and Y be two topological spaces and let $f, g : X \to Y$ be two continuous maps. A **homotopy** between f and g is a continuous map $H : [0,1] \times X \to Y$ such that $H(0,x) = f(x)$ and $H(1,x) = g(x)$. In such cases, we say that f and g are **homotopic**.

Let $f : X \to Y$ and $g : Y \to X$ be continuous maps. Let id_X and id_Y be the identity maps on X and Y. We say that X and Y are **homotopy equivalent** if the composition $g \circ f$ is homotopic to id_X and $f \circ g$ is homotopic to id_Y. We express this equivalence relation in symbols as $X \simeq Y$. The maps f and g are called **homotopy equivalences**.

Homotopy equivalence is a strictly stronger invariant than homology equivalence. In particular, homotopy equivalent spaces have the same homology. The converse does not hold.

When X is a subspace of Y, a **retraction** is a map $X \to Y$ that restricts to id_Y on Y. We say that Y is a **deformation retract** of X if there is a retraction homotopic to id_X. Intuitively, this means that we can realize a homotopy equivalence between X and Y by "shrinking" X to Y.

6.2.4 Singular Homology

It will occasionally be necessary to go beyond simplicial homology. **Singular homology** is a more general homology theory that allows us to define homology groups over topological spaces that are not simplicial complexes. The k-chains in singular homology are generated by all continuous maps from the standard k-simplex into the space. Thus, the dimension of the corresponding chain groups is uncountably infinite. We won't do computations on the singular homology, but we will use this theory at times to compare the homology of topological spaces that are not simplicial complexes. We will not give a formal treatment of the theory here, but instead refer the reader to the book by Hatcher [Hat01]. For now, we draw attention to two important facts regarding singular homology.

1. The singular homology of a simplicial complex is isomorphic to its simplicial homology [Hat01][Thm. 2.27].

2. The singular homology groups of homotopy equivalent spaces are isomorphic [Hat01][Cor. 2.11].

6.3 Persistent Homology

Filtrations. A **filtration** is a nested family of topological spaces, parameterized by single variable. We will focus on spaces that arise as subsets of Euclidean space. For example, a filtration $\{F^\alpha\}_{\alpha \geq 0}$ is a family of subsets $F^\alpha \subset \mathbb{R}^d$ such that $F^\alpha \subseteq F^\beta$ whenever $\alpha \leq \beta$. For any fixed α, the set F^α has the subspace topology inherited from $\mathbb{R}^d$.

One way to build a filtration is to consider a function $f : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ and define the sets in the filtration as

$$F^\alpha = \{x \in \mathbb{R}^d : f(x) \leq \alpha\} = f^{-1}[0, \alpha].$$

Such a filtration is called the **sublevel filtration** of the function f. If the homology groups of every space in a filtration have finite rank, then the filtration is said to be **tame**. All of the filtrations considered here will be tame.

One simple example of a sublevel filtration that arises often in both theory and practice is induced by the distance function to a finite point set $P \subset \mathbb{R}^d$. In this case, the distance function

to P is defined as

$$\mathbf{d}_P(x) = \min_{p \in P} |x - p|$$

and the α-subevel is known as the α**-offsets**

$$P^\alpha = \mathbf{d}_P^{-1}[0, \alpha] = \bigcup_{p \in P} \mathbf{ball}(p, \alpha).$$

Figure 6.3 shows P^α for several values of α.

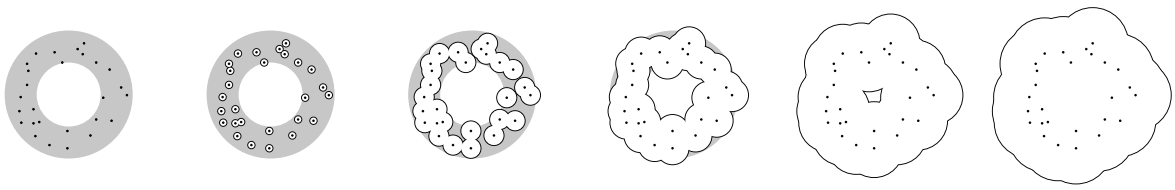

Figure 6.3: A set of points P is sampled from the gray annulus. The topology of the α-offsets, P^α, changes as the radius α increases.

The family of sublevel sets of $\mathbf{d}_P$ is called the **offsets filtration** of P. This filtration has played an important role in topological inference from point cloud data, where it has been used as a central theoretical tool for proving the correctness of existing techniques [CL07, CO08, CSEH05, NSW08a]. It is also known as the alpha-shape filtration [Ede95] and was the original filtration used for persistent homology [ELZ02].

We will also consider filtrations defined on simplicial complexes. We say that simplicial complex $\mathcal{K}$ is **filtered** by a function $f : \mathcal{K} \to \mathbb{R}_{\geq 0}$ if $\mathcal{K}^\alpha = f^{-1}[0, \alpha]$ is a simplicial complex for all $\alpha \in \mathbb{R}_{\geq 0}$.

Persistent Homology. Beginning with the work of Edelsbrunner, Letscher, and Zomorodian [ELZ02], persistent homology has expanded into a large and active research area. They showed that it is possible to track the the changes in homology over the course of a filtration. Instead of a static snapshot of the topology of F^α, we get a summary of all topological changes in $\{F^\alpha\}_{\alpha \geq 0}$, where α is interpreted as time. The inclusion map $i : F^\alpha \hookrightarrow F^\beta$ induces a homomorphism between the corresponding homology groups $i_k^{\alpha,\beta} : H_k(F^\alpha) \to H_k(F^\beta)$. A non-bounding cycle C is **born** at time α if α is the minimum such that C appears in F^α. A non-bounding cycle C **dies** at time β if $i_k^{\alpha,\beta}(C) = 0$ and $i_k^{\alpha,\gamma}(C) \neq 0$ for all $\gamma < \beta$. This is illustrated in Figure 6.4.

The output of the persistence algorithm is a collection of homology classes of each dimension along with their birth and death times in the filtration. This is the **persistence module** [ZC05]. One output representations is the **persistence diagram** that marks each homology class with a point in the plane, using the birth and death times as the x and y coordinates. Features that persist for a long time, those with a large gap between birth and death times, appear far from the diagonal $y = x$, whereas short-lived features (topological noise) concentrate around this diagonal.

An alternative representation of the output is a **persistence barcode**, which represents each

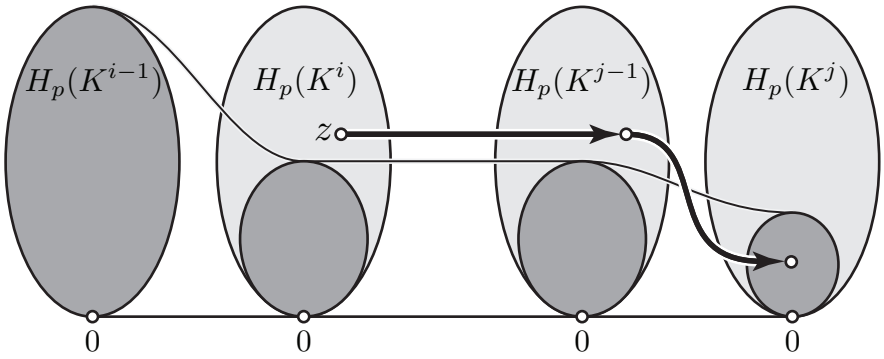

Figure 6.4: The inclusion map induces homomorphisms between the homology groups of the spaces in a filtration. The cycle z is born at time i and dies at time j.

feature as an interval using the birth and death times as left and right endpoints. Thus, features that persist for a long time, appear as long line segments, whereas short-lived features form short segments.

The Persistence Algorithm. The persistence diagram of a filtered simplicial complex can be computed using the persistence algorithm [ELZ02, ZC05]. Recall that for computing the homology of a simplicial complex, it suffices to compute the ranks of the boundary operators. This is easily done by row reduction.

For persistent homology a similar approach via row reduction works with just a little more care. We represent boundary matrices of the filtration so that the rows and columns are ordered according to the order of birth of the simplices. The persistence algorithm works by doing column reduction with the added constraint that we must work from left to right and only ever subtract columns to the left of the column we are trying to reduce. In the reduced matrix, the lowest one in any row (if there is one) corresponds to the birth and death of a topological feature.

If the total number of simplices is N, then the running time of the algorithm is $O(N^3)$. In theory, this may be improved to $O(N^\omega)$, where ω is the matrix multiplication constant. For this reason, computing simplicial filtrations of bounded size represents a large win for computing persistent homology.

Related Algorithmic Results for Persistence If a space is filtered by more than one parameter, we have a **multifiltration**. Carlsson, Singh, and Zomorodian gave a polynomial-time algorithm for multidimensional persistence [CSZ10]. An alternative approach for dealing with more than one parameter was given by Cohen-steiner, Edelsbrunner, and Morozov [CSEM06]. They give an algorithm for computing a sequence of persistence diagrams called a vineyard with the corresponding relationships between the features in the diagrams. Note that the vineyard approach differs from the multi-dimensional persistence. Their algorithm computes each successive diagram using a constant number of row operations for each pair of simplices transposed between the two filtrations. Thus, given a complex $\mathcal{K}$ with n simplices and two different filtrations on $\mathcal{K}$, the persistence diagram of the

second filtration can be computed from the persistence diagram of the first filtration in $O(n^2)$ time in the worst case, with better performance expected if the filtrations are similar. Zigzag persistence is generalization of persistent homology developed by Carlsson and de Silva that allows for "filtrations" that are do not necessarily grow monotonically [CdSM09, CdS10]. Milosavljevic, Morozov, Skraba gave an algorithm that computes Zigzag persistence for simplicial complexes with deletions and insertions in $O(n^\omega + n^2 \log n)$ time where ω is the matrix multiplication constant [MMS11].

6.4 Persistence and Geometry

Nerves. Given a collection of closed sets U, the **nerve** of U is a simplicial complex with vertex set U and simplices for each collection of sets of U that have a common intersection. The Nerve Theorem states that the nerve of collection of sets is homotopy equivalent to their union if all intersections of finitely many sets are either empty or contractible. The family U is called a **closed cover** of $\bigcup_{u \in U} u$. If the sets of U are all convex, then they easily satisfy the hypothesis of the Nerve Theorem and we call U a **good closed cover**.

Theorem 6.4.1. *If X is a topological space and $\mathcal{U} = U_1, \ldots, U_k$ is a good closed cover, then X and the nerve $\mathcal{NU}$ are homotopy equivalent.*

Since homotopy equivalences induce isomorphisms between homology groups, we may interpret the Nerve Theorem as a way to translate between the homology of general spaces and the simplicial homology of some discretization. In order to extend this equivalence to persistent homology requires a little more work. First, we recall the Persistence Equivalence Theorem of Zomorodian and Carlsson [ZC05], which gives a general condition for two filtrations to have the same persistent homology. Then, we will state the Persistent Nerve Lemma of Chazal and Oudot [CO08] that gives proves sufficient conditions for a filtration and a filtration of good closed covers to have the same persistent homology.

Theorem 6.4.2 (Persistence Equivalence Theorem [ZC05])**.** *Let $\{F_\alpha\}$ and $\{G_\alpha\}$ be filtrations and let $\phi_\alpha : H(F_\alpha) \to H(G_\alpha)$ be an isomorphism for all $\alpha \geq 0$. Consider the commutative diagram where the maps $H(F_\alpha) \to H(F_\alpha)$ and $H(G_\alpha) \to H(G_\alpha)$ are induced by the inclusions $F_\alpha \hookrightarrow F_\beta$ and $G_\alpha \hookrightarrow G_\beta$:*

$$\begin{array}{ccc} H_k(F_\alpha) & \to & H_k(F_\beta) \\ \phi_\alpha \downarrow & & \downarrow \phi_\beta \\ H_k(G_\alpha) & \to & H_k(G_\beta) \end{array}$$

If the diagram commutes for all $0 \leq \alpha \leq \beta$, then $\{F_\alpha\}$ and $\{G_\alpha\}$ have the same persistence diagram.

The Persistent Nerve Lemma says that if we have a good closed cover of a space in a filtration, and the sets in the covers themselves form filtrations, then the filtration and its nerve filtration have the same persistent homology. This is a very handy tool as it allows us to move easily between filtrations on geometric spaces and filtered simplicial complexes.

Lemma 6.4.3 (The Persistent Nerve Lemma [CO08])**.** *Let $X \subseteq X'$ be two paracompact spaces, and let $\mathcal{U} = \{U_a\}_{a\in A}$ and $\mathcal{U}' = \{U'_{a'}\}_{a'\in A'}$ be good closed covers of X and X' respectively, based on finite parameter sets $A \subseteq A'$ such that $U_a \subseteq U'_a$ for all $a \in A$. Then, the homotopy equivalences $\mathcal{NU} \to X$ and $\mathcal{NU}' \to X'$ provided by the Nerve Theorem [Hat01, Corollary 4G.3] commute with the canonical inclusions $X \hookrightarrow X'$ and $\mathcal{NU} \hookrightarrow \mathcal{NU}'$ at homology level.*

Equivalence of Filtrations. It is often useful to replace one filtration with another that has an equivalent persistence diagram. The following lemma gives three sufficient conditions for two filtrations to yield equivalent persistence diagrams.

Lemma 6.4.4. *Given two filtrations $\mathcal{F} = \{F_\alpha\}_{\alpha\geq c}$ and $\mathcal{G} = \{G_\alpha\}_{\alpha\geq c}$, $\mathcal{F}$ and $\mathcal{G}$ have identical persistence diagrams if any of the following conditions are met:*

1. *The canonical inclusion $F_\alpha \hookrightarrow G_\alpha$ is a homotopy equivalence for all $\alpha \geq c$.*
2. *$\mathcal{F}$ and $\mathcal{G}$ are filtered simplicial complexes and there is a bijection $\phi : F_\infty \to G_\infty$ that restricts to a bijection for each F^α.*
3. *There exists good closed covers $\mathcal{U}_\alpha = \{U_i^\alpha\}_{i\in A}$ of F_α and $\mathcal{U}'_\alpha = \{U_i'^\alpha\}_{i\in A}$ of F'_α satisfying the conditions of the Persistent Nerve Lemma for all $\alpha' \geq \alpha \geq c$ and G_α (G'_α) is the nerve of $\mathcal{U}_\alpha$ ($\mathcal{U}'_\alpha$ respectively).*

Proof. In all three cases, the assumption leads to the existence of isomorphisms between the homology groups $H_k(F_\alpha)$ and $H_k(G_\alpha)$ for all k and all $\alpha \geq c$. By the Persistence Equivalence Theorem, it suffices to prove that these isomorphisms commute with the homomorphisms induced by the inclusions $F_\alpha \hookrightarrow F_\beta$ and $G_\alpha \hookrightarrow G_\beta$ for all $c \leq \alpha \leq \beta$. That is, we need to show that the following diagram commutes.

$$\begin{array}{ccc} H_k(F_\alpha) & \to & H_k(F_\beta) \\ \cong\downarrow & & \downarrow\cong \\ H_k(G_\alpha) & \to & H_k(G_\beta) \end{array}$$

The first condition and second conditions suffice because all of the maps commute at the set level and therefore at the homology level as well. The third condition is a direct application of the Persistent Nerve Lemma. □

Notations for Complexes and Filtrations. We adopt the following notational conventions for filtered complexes. Variables in uppercase calligraphy font such as $\mathcal{A}$ are simplicial complexes. Superscripts on a complex indicate a subcomplex from the filtration corresponding to a particular sublevel or scale, i.e. $\mathcal{A}^\alpha$ is the subcomplex of $\mathcal{A}$ at scale α. A filtration in one variable will be denoted $\{\mathcal{A}_k^\alpha\}$ where it is assumed that the parameter is α ranging over all nonnegative real numbers and k is fixed. For multifiltrations subscripts indicate the value of the second parameter, k. The only multifiltrations we will use come from sequences of distance functions parameterized by a positive integer k. Such a multifiltration will be denoted $\{\mathcal{A}_k^\alpha\}_{k,\alpha}$ where it is assumed that α ranges over the nonnegative reals and k ranges over the positive integers.

6.5 Approximating Persistence Diagrams

The notion of approximate persistence diagrams comes naturally out of the notion of stability of persistence diagrams [CSEH05]. For stability, the goal is to show that two similar inputs yield similar outputs. For approximation, we replace one of these inputs with the true filtration that we want to approximate, P_k^α in our case, and argue that our approximate filtration will produce a persistence diagram that is provably close to the persistence diagram of the true of filtration.

First we need a notion of closeness for persistence diagrams. The bottleneck distance between two multisets $A, B \subset [0, +\infty]^2$ is defined as

$$\mathbf{d}_B^\infty(A, B) = \min_\gamma \max_{p \in A} \|p - \gamma(p)\|_\infty,$$

where $\|\cdot\|_\infty$ denotes the l_∞-norm and γ ranges over all bijections from A to B. To make sure that such bijections always exist, the diagonal $\{(x, x) : x \in [0, +\infty]\}$ is added to every persistence diagram infinite multiplicity.

Proximity between filtrations is defined in terms of mutual nesting: specifically, two tame filtrations $\mathcal{F}, \mathcal{G}$ are said to be **ε-interleaved** if we have $F_\alpha \subseteq G_{\alpha+\varepsilon}$ and $G_\alpha \subseteq F_{\alpha+\varepsilon}$ for all $\alpha \geq 0$. Under this condition, it is known that the persistence diagrams $\mathrm{D}\mathcal{F}$ and $\mathrm{D}\mathcal{G}$ are ε-close in the bottleneck distance [CCSG+09, CSEH05]. The formal statement goes as follows:

Theorem 6.5.1 (Stability [CCSG+09, CSEH05])**.** *If two tame filtrations $\mathcal{F}, \mathcal{G}$ are ε-interleaved, then $\mathbf{d}_B^\infty(\mathrm{D}\mathcal{F}, \mathrm{D}\mathcal{G}) \leq \varepsilon$.*

Multiplicative interleaving. Two filtrations $\mathcal{F}, \mathcal{G}$ are **multiplicatively c-interleaved** if

$$F_{\frac{\alpha}{c}} \subseteq G_\alpha \subseteq F_{c\alpha}$$

for all $\alpha \geq 0$. For a filtration $\{F_\alpha\}$, we can reparameterize it on the natural logarithmic scale by defining

$$F_\alpha^{\ln} = F_{e^\alpha},$$

and

$$\ln \mathcal{F} := \{F_\alpha^{\ln}\}_{\alpha \in \mathbb{R}}$$

Multiplicative c-interleaving of $\mathcal{F}$ and $\mathcal{G}$ implies additive $\ln c$-interleaving of $\ln \mathcal{F}$ and $\ln \mathcal{G}$ As a result, multiplicative interleaving of filtrations implies the following weaker form of proximity between their persistence diagrams, where the notation $\mathbf{d}_{\mathrm{D}}^{\ln}(\mathcal{F}, \mathcal{G})$ (called **log-diagram distance**) stands for the quantity $\mathbf{d}_B^\infty(\mathrm{D} \ln \mathcal{F}, \mathrm{D} \ln \mathcal{G})$:

Corollary 6.5.2. *If two filtrations $\mathcal{F}, \mathcal{G}$ are multiplicatively c-interleaved, then $\mathbf{d}_{\mathrm{D}}^{\ln}(\mathcal{F}, \mathcal{G}) \leq \ln c$.*

A persistence diagram $\mathrm{D}\mathcal{F}$ is **$(1+\varepsilon)$-approximation** to a persistence diagram $\mathrm{D}\mathcal{G}$ if $\mathbf{d}_{\mathrm{D}}^{\ln}(\mathcal{F}, \mathcal{G}) \leq \ln(1+\varepsilon)$. This means that every point in the first diagram can be matched to a point in the second such that the birth and death times differ by at most a factor of $1 + \varepsilon$.

Multiplicative approximation makes sense in the geometric setting because it corresponds to a noise model in which the scale of the noise depends on the scale of the features. Although tighter guarantees are possible with an additive interleaving, they come at the cost of assuming an absolute bound on the scale of the noise everywhere. We can thus state a special case of the Strong Stability Theorem of Chazal et al. [CCSG$^+$09], rephrased into the language of multiplicative approximations.

Theorem 6.5.3. *Let $\{F^\alpha\}$ and $\{G^\alpha\}$ be two tame filtrations. If $F^{\alpha/c} \subseteq G^\alpha \subseteq F^{c\alpha}$ for all $\alpha \geq 0$, then the persistence diagram of $\{F^\alpha\}$ is a c-approximation to the persistence diagram of $\{G^\alpha\}$.*

Chapter 7

Geometric Persistent Homology for Offsets using Meshes

7.1 Overview

We apply ideas from mesh generation to approximate the persistence diagram of the full offsets filtration of a point cloud $P \subset \mathbb{R}^d$. Classical approaches rely on the Čech, Rips, α-complex, or witness complex filtrations of P, whose complexities scale badly with d. For instance, the α-complex filtration incurs the $n^{\Omega(d)}$ size of the Delaunay triangulation, where $n = |P|$. The common alternative is to truncate the filtrations when the sizes of the complexes become prohibitive, possibly before discovering the most relevant topological features. In this chapter we present a new collection of filtrations based on the Delaunay triangulation of a mesh, whose sizes are reduced to $2^{O(d^2)}n$. A nice property of these filtrations is is that they are interleaved multiplicatively with the family of offsets of P. Thus, the persistence diagram of P can be approximated in $2^{O(d^2)}n^3$ time in theory, with a near-linear observed running time in practice. Consequently, our approach remains tractable in medium dimensions, say 4 to 10.

7.2 Point sets at different scales

Persistent homology is a powerful tool for understanding the topological structure of a point cloud across different scales. Given an appropriate simplicial complex and filtration, the short-lived features are sampling noise, while long-lived features are significant. The persistence diagram aids in analyzing the shape from which a point cloud was drawn.

Several filtrations have been used with success in the past, including the α-complex [Ede95] and witness complex [dS08, dSC04] filtrations, which are based on the Delaunay triangulation of P or an approximation of it, and the Čech [ES52] and Vietoris-Rips [Vie27] filtrations, which are derived from the nerves of collections of congruent balls centered at the data points. In practice however, the cost to build them makes their use prohibitive, even in medium dimensions, say 4 to 10. When

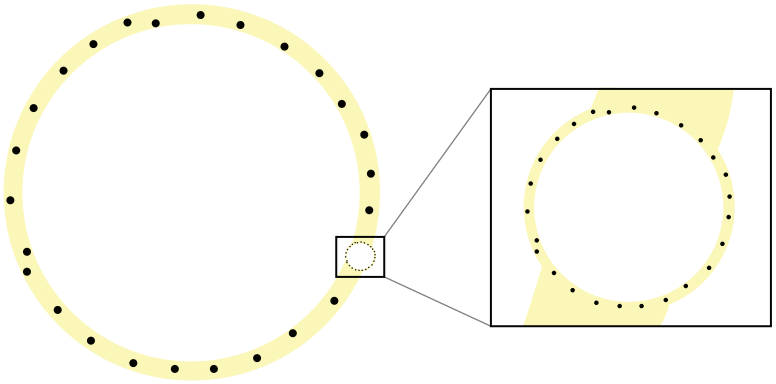

Figure 7.1: When topological features appear at dramatically different scales, classic filtrations reach a very high complexity before the largest features can be captured.

α becomes large, the size of the α-complex approaches that of the Delaunay triangulation: $n^{\Omega(d)}$ in d dimensions even for some relatively "nice" inputs [Eri01]. The sizes of the Čech, Rips and (relaxed) witness complexes grow even more quickly, as $2^{\Omega(n)}$. If one restricts the complex to only include simplices up to dimension d, the complexity is still $n^{\Theta(d)}$.

To avoid this difficulty, researchers usually resort to truncating the filtrations at a prescribed size limit. Truncation is equivalent to looking at the data at small scales only, and can make the algorithm miss relevant structures at larger-scales. This can happen even in simple scenarios, such as the one depicted in Figure 7.1. Another example of interest, inspired from [GO07], is described in Figure 7.2 (left): it consists of a point cloud sampled evenly from a helicoidal curve drawn on the Clifford torus in $\mathbb{R}^4$. In this case, the point cloud admits at least three candidate underlying spaces: at a small scale, the curve; at a larger scale, the torus; and at an even larger scale, the 3-sphere of radius $\sqrt{2}$ on which the Clifford torus is naturally embedded. One might also add the point cloud itself and $\mathbb{R}^4$ at either ends of the spectrum.

In order to analyze such data sets at different scales using only truncated filtrations, Chazal and Oudot [CO08] proposed a **landmarking** strategy in the spirit of [GO07], which maintains a growing subset of the data points, on which the simplicial complexes are built. However, their approach produces a weaker form of data representation than persistence diagrams, which does not explicitly correlate the features visible at different scales. As a result, they can get false positives when retrieving the set of persistent topological features.

How Voronoi refinement helps. We generate a mesh $M = P \cup S$ of well-spaced points and order the simplices of its Delaunay triangulation according to a filter $t : \mathrm{Del}_M \to \mathbb{R}$. Several different filters are analyzed, yielding filtrations with different properties: some are easier to build, others come with better approximation guarantees. The choice of a particular filter depends on the application and is left to the user. All of these filters are based on distances to the input point cloud P. We show that the corresponding filtrations are interleaved on a logarithmic scale with the filtration of the offsets of P, and thus produce accurate approximate persistence diagrams. Computing the

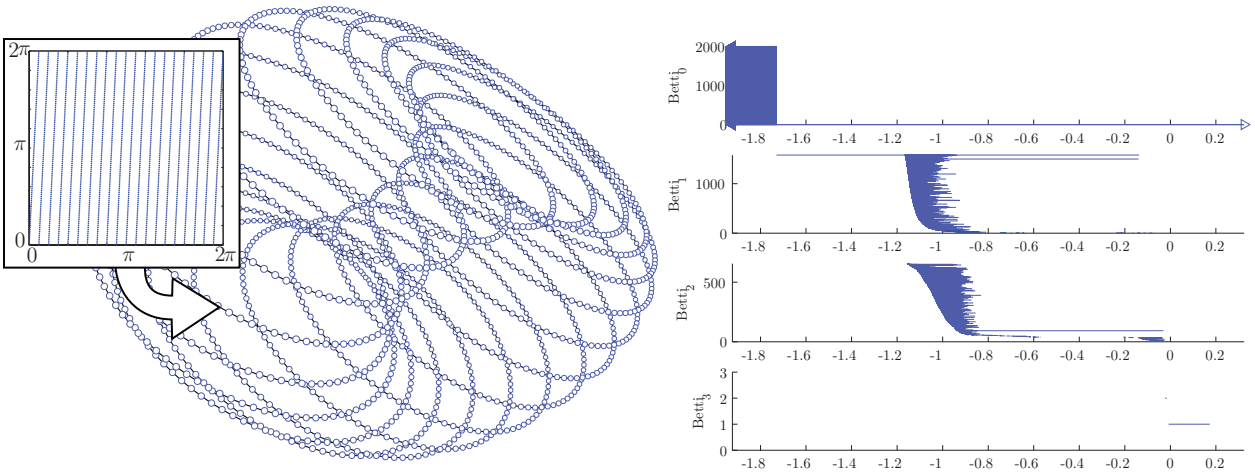

Figure 7.2: The Clifford data set. Left: point cloud sampled uniformly along a periodic curve in $[0, 2\pi]^2$, then mapped onto a helicoidal curve drawn on the Clifford torus in $\mathbb{R}^4$ via the canonical embedding $(u, v) \mapsto (\cos u, \sin u, \cos v, \sin v)$. The rendering is via a projection into $\mathbb{R}^3$. Right: log-scale barcode obtained on this data set using the filtration of Section 7.4.2.

persistence diagram takes time cubic in the number of simplices and thus dominates our worst-case $2^{O(d^2)}n^3$ overall runtime. This bound, though large, is still a significant improvement over $n^{\Omega(d)}$. Moreover, in practice, the persistence diagram computation takes near-linear time (on an input with $2^{O(d^2)}n$ simplices), which makes our approach tractable in small to medium dimensions (4-10) for moderate input sizes (thousands to tens of thousands of points). A preliminary implementation bears out these predictions (see Section 7.7).

We first present a simplified version in Section 7.4 that produces a filtration that is $\log(\tau)$-interleaved with the offsets filtration of P, for some constant $\tau \geq 2$. The size of this filtration is $2^{O(d^2)}n \log(\Delta)$, where Δ denotes the spread of P. We then show in Section 7.5 how the interleaving between our filtration and the offsets filtration can be tightened; we produce persistence diagrams that are accurate within any arbitrarily small error. Finally, in Section 7.6 we concentrate on the size of the filtration and show how to eliminate its dependence on the spread using hierarchical meshes.

7.3 Preliminaries

7.3.1 Clipped Voronoi Diagrams and Voronoi Refinement

Let P be a finite set of points in general position in $\mathbb{R}^d$ with Voronoi diagram Vor_P and Delaunay triangulation Del_P. Since P is in general position, Del_P is an embedded simplicial complex in $\mathbb{R}^d$,

whose underlying space is $\text{conv}(P)$.

Given a compact, convex bounding domain BB containing P, we consider the restrictions of the Voronoi diagram and Delaunay triangulation to BB. Specifically, given a point $p \in P$, we call $\text{Vor}_\square(p)$ its **Voronoi cell clipped to** BB:

$$\text{Vor}_\square(p) = \text{Vor}_P(p) \cap BB.$$

We call $\text{Vor}_\square(P)$ the **Voronoi diagram clipped to** BB, and $\text{Del}_\square(P)$ is its dual complex, which is a subcomplex of Del_P. The aspect ratio of a clipped Voronoi cell is the same as the aspect ratio of the unclipped cell as long as the clipping does not eliminate any corners of a Voronoi cell. In particular, this happens when the clipping just eliminates the unbounded portion of a Voronoi cell on on the boundary. By constructing a bounding domain and boundary net as in Section 3.5, we can assume that no cell has its aspect ratio changed by the clipping.

Voronoi Refinement. The NetMesh algorithm takes a finite point cloud P as input and returns a finite superset M of P that satisfies the following properties:

(i) M is a point sampling of some bounding ball BB of radius $O(\textbf{diameter}(P))$ around the input point cloud P,
(ii) The Delaunay triangulation clipped to BB, $\text{Del}_\square(M)$, is equal to the full Delaunay triangulation Del_M,
(iii) The aspect ratios of the clipped Voronoi cells of the points of M are bounded from above by an absolute constant $\tau \geq 2$,
(iv) The complexity of Del_M is $2^{O(d^2)}|M|$,
(v) The size of M is $2^{O(d)} n \log \Delta$, where Δ is the **spread** of P.

The NetMesh algorithm can produce M in near-optimal $2^{O(d^2)} n \log n + |M|$ time. As shown in Chapter 4, it is possible to reduce the output-sensitive term $|M|$ to $2^{O(d^2)} n$ by considering only subsets that admit a well-paced ordering. This technique will be used in Section 7.6 to eliminate the dependence on the spread in $|M|$.

7.4 The α-mesh filtration

Our strategy is to build a mesh M on the input set P, and then filter Del_M to obtain a filtration that can be related to the sublevel filtration of $\mathbf{d}_P$. In Section 7.4.1 we present a simplified version of the filter. The analysis of the basic filter relies on the same key ingredients as the full version and leads to a partial approximation result (Theorem 7.4.5). In Section 7.4.2 we explain the limitations of the basic filter and the modifications required to obtain a full approximation guarantee (Theorem 7.4.8).

7.4.1 Basic filter

Our input is a finite set P of points in general position in $\mathbb{R}^d$. We first apply the NetMesh algorithm to construct a superset $M \supseteq P$ that satisfies conditions (i) through (v) of Section 7.3.1.

We then define the filter $t : \mathrm{Del}_M \to \mathbb{R}$ as follows.

$$t(\sigma) = \max_{v \in \sigma} d_P(v).$$

We define the α**-mesh filtration** $\{D_M^\alpha\}_{\alpha \geq 0}$ formally as the sublevel filtration of t. That is, for all $\alpha \geq 0$, we let D_M^α be the subcomplex of Del_M defined as

$$D_M^\alpha = \{\sigma \in \mathrm{Del}_M : t(\sigma) \leq \alpha\}.$$

Equivalently, D_M^α contains all simplices of Del_M whose vertices are contained in P^α. Note that if σ' is a face of σ then $t(\sigma') \leq t(\sigma)$, so the spaces forming the filtration are proper simplicial complexes, and we have $D_M^\alpha \subseteq D_M^\beta$ for all $0 \leq \alpha \leq \beta$.

Intuitively, the basic filter sorts the simplices of Del_M by their distance to P, simulating within Del_M the growth of the offsets of P (see Figure 7.3 (right) for an illustration). As will be shown in the analysis, the simulation process "works" because Voronoi cells have bounded aspect ratios.

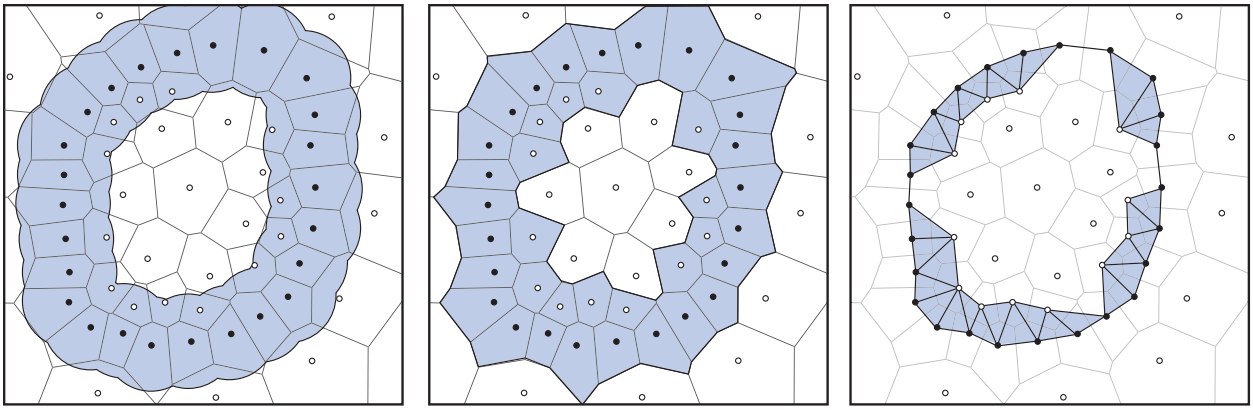

Figure 7.3: From left to right: the offset P^α, the α-Voronoi V_M^α and its dual α-mesh D_M^α.

Theoretical analysis. Our goal is to relate $\{D_M^\alpha\}_{\alpha \geq 0}$ to the offsets filtration $\{P^\alpha\}_{\alpha \geq 0}$. We do the analysis in terms of a dual filtration, $\{V_M^\alpha\}_{\alpha \geq 0}$, based on the clipped Voronoi diagram $\mathrm{Vor}_\square(M)$ (see Figure 7.3 (center) for an illustration). To each point $v \in M$ we assign a closed convex set $U_\alpha(v)$ as follows:

$$U_\alpha(v) = \begin{cases} \emptyset & \text{if } \alpha < t(v), \\ \mathrm{Vor}_\square(v) & \text{otherwise.} \end{cases} \tag{7.1}$$

The α-Voronoi filtration $\{V_M^\alpha\}_{\alpha \geq 0}$ is defined by

$$V_M^\alpha = \bigcup_{v \in M} U_\alpha(v).$$

The collection $\mathcal{U}_\alpha = \{U_\alpha(v)\}_{v\in M}$ forms a closed cover of V_M^α. Let $\mathcal{NU}_\alpha$ denote the nerve of this cover. Both D_M^α and $\mathcal{NU}_\alpha$ are embedded as subcomplexes of the full simplex 2^M over the vertex set M, and the following lemma stresses their relationship:

Lemma 7.4.1. *For all $\alpha \geq 0$, the subcomplexes D_M^α and $\mathcal{NU}_\alpha$ of the full simplex 2^M are equal.*

Proof. Consider first the case of a 0-dimensional simplex $\sigma = \{v\}$. The definition of D_M^α states that $\{v\} \in D_M^\alpha$ if and only if $\mathbf{d}_P(v) \leq \alpha$, which is also the criterion for which U_α is nonempty and hence belongs to the collection $\mathcal{U}_\alpha$. Thus, $\{v\} \in D_M^\alpha \Leftrightarrow \{v\} \in \mathcal{NU}_\alpha$.

Consider now the case of a k-simplex $\sigma = \{v_0, \cdots, v_k\}$ for $k > 0$. The definition of D_M^α states that $\sigma \in D_M^\alpha$ if and only if $\sigma \in \mathrm{Del}_M$ and $\max_i t(v_i) \leq \alpha$, which is equivalent to $\bigcap_{i=0}^k \mathrm{Vor}(v_i) \neq \emptyset$ and $\max_i t(v_i) \leq \alpha$. Since $\mathrm{Del}_s q(M) = \mathrm{Del}_M$ (assertion (ii) of Section 7.3.1),

$$\bigcap_{i=0}^k \mathrm{Vor}(v_i) \neq \emptyset \Leftrightarrow \bigcap_{i=0}^k \mathrm{Vor}(v_i) \cap BB \neq \emptyset \Leftrightarrow \bigcap_{i=0}^k \mathrm{Vor}_\square(v_i) \neq \emptyset.$$

Hence, $\sigma \in D_M^\alpha$ if and only if $\sigma \in \mathcal{NU}_\alpha$. □

Since the sets U_α in the cover of V_M^α are convex, we can apply the Persistent Nerve Lemma and Lemma 7.4.1 to relate the persistence diagrams of $\{V_M^\alpha\}_{\alpha\geq 0}$ and $\{D_M^\alpha\}_{\alpha\geq 0}$.

Lemma 7.4.2. *For all $r \geq 0$, the persistence diagrams of the filtrations $\{V_M^\alpha\}_{\alpha\geq r}$ and $\{D_M^\alpha\}_{\alpha\geq r}$ are identical.*

Proof. Fix any $r \geq 0$. We first use Lemma 6.4.4 to show that $\{V_M^\alpha\}$ and $\{\mathcal{NU}_\alpha\}$ have identical persistence diagrams. To apply Lemma 6.4.4 it suffices to observe that the sets $U_\alpha(v)$ satisfy the conditions of the Persistent Nerve Lemma:

- the sets $U_\alpha(v)$ cover V_M^α,
- each $U_\alpha(v)$ is are all convex, and thus form a good closed cover, and
- for $\alpha < \beta$ $U_\alpha(v) \subseteq U_\beta(v)$.

Lemma 7.4.1 implies that $\{D_M^\alpha\}_{\alpha\geq r}$ and $\{\mathcal{NU}_\alpha\}_{\alpha\geq r}$ have identical persistence diagram. So, the persistence diagrams of the filtrations $\{V_M^\alpha\}_{\alpha\geq r}$ and $\{D_M^\alpha\}_{\alpha\geq r}$ are also identical. □

Let the clipped offsets be defined as follows, in analogy with the clipped Voronoi cells:

$$P_\square^\alpha = \{x \in BB : \mathbf{d}_P(x) \leq \alpha\}.$$

Let also $r_P = \frac{1}{2}\max_{p\in P} \mathbf{d}_{M\setminus\{p\}}(p)$.

Lemma 7.4.3. *For all $\alpha \geq r_P$, $V_M^{\alpha/\tau} \subseteq P_\square^\alpha \subseteq V_M^{\tau\alpha}$.*

Proof. Let x be a point of $V_M^{\alpha/\tau} \subseteq BB$, and let $v \in M$ be such that $x \in U_{\alpha/\tau}(v)$. Let also $p \in P$ be closest to v. If $v \in S$, then the fact that $U_{\alpha/\tau}(v) \neq \emptyset$ implies that $|v - p| \leq \alpha/\tau$ and

$U_{\alpha/\tau}(v) = \mathrm{Vor}_\square(v)$. This implies that $|x-p| \leq |x-v| + |v-p| \leq |x-v| + \alpha/\tau$. Now, the aspect ratio of $\mathrm{Vor}_\square(v)$ is at most τ (assertion (iii) of Section 7.3.1), implying that

$$|x-v| \leq \frac{\tau}{2}\mathbf{d}_{M\setminus\{v\}}(v) \leq \frac{\tau}{2}|v-p| \leq \frac{\alpha}{2}.$$

Thus, $\mathbf{d}_P(x) \leq |x-p| \leq \alpha(\frac{1}{2}+\frac{1}{\tau})$. Since $\tau \geq 2$, we conclude that $\mathbf{d}_P(x) \leq \alpha$. If now $v \in P$, then the aspect ratio condition (assertion (iii) of Section 7.3.1) implies that $|x-v| \leq \frac{\tau}{2}\mathbf{d}_{M\setminus\{v\}}(v) \leq r_P \leq \alpha$. Hence, in all cases we have $\mathbf{d}_P(x) \leq |x-v| \leq \alpha$, which means that $x \in P^\alpha_\square$.

Let now x be a point of $P^\alpha_\square$, and let $v \in M$ and $p \in P$ be closest to x. Then, x belongs both to $\mathrm{Vor}_\square(v)$ and to the Euclidean ball of center p and radius α. It follows that $\mathbf{d}_P(v) \leq |v-p| \leq |v-x| + |x-p| \leq 2|x-p| \leq 2\alpha \leq \tau\alpha$. This means that $U_{\tau\alpha}(v) = \mathrm{Vor}_\square(v)$, and since $x \in \mathrm{Vor}_\square(v)$, $x \in V_M^{\tau\alpha}$ as desired. □

We now show that clipping offsets to BB does not change homotopy type as long as BB contains the points set. Specifically, the following lemma shows that $P^\alpha_\square \hookrightarrow P^\alpha$ is a homotopy equivalence. A proof may be found in Section 7.9.

Lemma 7.4.4. *Let B be a compact, convex set. If $Q \subset B$ then for all $\alpha \geq 0$, the canonical inclusion $Q^\alpha \cap B \hookrightarrow Q^\alpha$ is a homotopy equivalence.*

Using the above results we can conclude our analysis, which relates the diagrams of the **truncated** filtrations $\{P^\alpha\}_{\alpha \geq r_P}$ and $\{D^\alpha_M\}_{\alpha \geq r_P}$:

Theorem 7.4.5. *On the natural logarithmic scale, the persistence diagrams of $\{P^\alpha\}_{\alpha \geq r_P}$ and $\{D^\alpha_M\}_{\alpha \geq r_P}$ are $\ln\tau$-close in bottleneck distance, i.e. $\mathbf{d}_{\mathrm{D}}^{\ln}(\{P^\alpha\}_{\alpha \geq r_P},\ \{D^\alpha_M\}_{\alpha \geq r_P}) \leq \ln\tau$.*

Proof. By Lemma 7.4.4, the canonical inclusions $P^\alpha_\square \hookrightarrow P^\alpha$ and $P^\beta_\square \hookrightarrow P^\beta$ are homotopy equivalences that commute with the inclusions $P^\alpha_\square \hookrightarrow P^\beta_\square$ and $P^\alpha \to P^\beta$ for all $\beta \geq \alpha \geq 0$, so the filtrations $\{P^\alpha\}_{\alpha\geq 0}$ and $\{P^\alpha_\square\}_{\alpha\geq 0}$ have identical persistence diagrams. In addition, Lemma 7.4.2 implies that $\{D^\alpha_M\}_{\alpha\geq 0}$ and $\{V^\alpha_M\}_{\alpha\geq 0}$ have identical persistence diagrams. The result follows then from the interleaving of the truncated filtrations $\{P^\alpha_\square\}_{\alpha\geq r_P}$ and $\{V^\alpha_M\}_{\alpha\geq r_P}$ (Lemma 7.4.3) and its consequences on the proximity of their persistence diagrams (Corollary 6.5.2). □

Intuitively, Theorem 7.4.5 means that homological features appearing in the offsets filtration after time $\alpha = r_P$ are captured by the α-mesh filtration with approximately same birth and death times on the natural logarithmic scale. Features appearing before r_P and dying after r_P are also captured, but starting at times as late as r_P (the death times remaining approximately the same). Finally, features appearing and dying before r_p may not be captured at all.

7.4.2 Complete filter

The obvious drawback of the basic filter is that it only enables us to approximate the persistence diagram of the offsets filtration after a certain time (Theorem 7.4.5). The reason is clear from the

proof of Lemma 7.4.3: even though we have $P^\alpha_\square \subseteq V^{\tau\alpha}_M$ for all $\alpha \geq 0$, the symmetric inclusion $V^\alpha_M \subseteq P^{\tau\alpha}_\square$ only holds when $\alpha \geq r_P$, since the clipped Voronoi cells of the input points appear in V^α_M as soon as time $\alpha = 0$ and they are not covered by P^α before $\alpha = \tau r_P$. In the dual α-mesh, this phenomenon translates into the appearance of edges between the points of P as early as time $\alpha = 0$, whereas such edges should normally appear when α-balls around these points touch one another. In this section we propose a solution to this issue, which consists in modifying the filter of Del_M so as to somewhat delay the appearances of the simplices incident to the points of P in the α-mesh filtration. The rest of the approach remains unchanged, namely: we run the NetMesh algorithm on the input point cloud P to get our vertex set $M \supseteq P$, then we define a modified filter $\tilde{t} : \mathrm{Del}_M \to \mathbb{R}$ and build its sublevel filtration $\{\tilde{D}^\alpha_M\}_{\alpha\geq 0}$.

Filter modification. We follow the above recommendations and modify our filter to be a little more careful in the region close to the input points. Let $s(v) = \frac{1}{2} f_M(v)$ if $v \in P$ and $s(v) = \mathbf{d}_P(v)$ otherwise, where f_M denotes the Ruppert local feature size of M. Then, we modify our filter as follows:

$$\tilde{t}(\sigma) = \begin{cases} d_P(v) & \text{if } \sigma \text{ is a vertex } v \in M, \\ \max\limits_{v\in\sigma} s(v) & \text{otherwise.} \end{cases}$$

The difference between this filter and the one of Section 7.4.1 resides in the second item, which delays the times at which the Delaunay simplices incident to the points of P appear.

Theoretical analysis. We redo the analysis of Section 7.4.1 using a modified dual filtration $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$. We only detail the changes to be made to the statements and proofs. Each point $v \in M$ is assigned a convex set $\tilde{U}_\alpha(v)$ as follows:

$$\tilde{U}_\alpha(v) = \begin{cases} \emptyset & \text{if } \alpha < \tilde{t}(v), \\ \overline{\mathbf{ball}}(v, \alpha) & \text{if } v \in P \text{ and } \tilde{t}(v) \leq \alpha < s(v), \\ \mathrm{Vor}_\square(v) & \text{otherwise.} \end{cases} \tag{7.2}$$

The filtration $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$ is defined by $\tilde{V}^\alpha_M = \bigcup_{v\in M} \tilde{U}_\alpha(v)$ for every $\alpha \geq 0$. As in Section 7.4.1, the collection of the sets $\tilde{U}_\alpha(v)$ forms a good closed cover of $\tilde{V}^\alpha_M$, and the sets themselves are monotonically increasing with α. Therefore, the same arguments as in the proof of Lemma 7.4.2 show that $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$ and $\{\tilde{\mathcal{NU}}_\alpha\}_{\alpha\geq 0}$ have identical persistence diagrams, where $\tilde{\mathcal{NU}}_\alpha$ denotes the nerve of the cover $\{\tilde{U}_\alpha(v) : v \in M,\ \tilde{U}_\alpha(v) \neq \emptyset\}$. Moreover, the following analog of Lemma 7.4.1 relates $\{\tilde{\mathcal{NU}}_\alpha\}_{\alpha\geq 0}$ to $\{\tilde{D}^\alpha_M\}_{\alpha\geq 0}$:

Lemma 7.4.6 (analog of Lemma 7.4.1)**.** *For all $\alpha \geq 0$, $\tilde{D}^\alpha_M$ and $\tilde{\mathcal{NU}}_\alpha$ (viewed as subcomplexes of the full simplex 2^M) are identical.*

Proof. Let $\sigma = \{v_0, \cdots, v_k\} \subseteq M$ be a simplex. If $k = 0$, then our definitions imply that the vertex $\sigma = \{v_0\}$ appears in $\tilde{\mathcal{NU}}_\alpha$ and in $\tilde{D}^\alpha_M$ at the same time $\tilde{t}(v_0)$. If $k > 0$, then it follows from Eq. (7.2) that σ must be a simplex of Del_M in order to appear in $\tilde{\mathcal{NU}}_\alpha$, since the balls $\mathbf{ball}(p, \alpha)$ are

pairwise disjoint and disjoint from the Voronoi cells of the other points. Now, σ appears in $\tilde{\mathcal{NU}}_\alpha$ at time $\max_{i=1,\cdots,k} s(v_i)$, which is also the time at which σ appears in $\tilde{D}^\alpha_M$. □

It follows from Lemma 7.4.6 and preceding discussion that the filtrations $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$ and $\{\tilde{D}^\alpha_M\}_{\alpha\geq 0}$ have identical persistence diagrams. Defining the clipped offsets $P^\alpha_\square$ as in Section 7.4.1, we can make the interleaving between $\{P^\alpha_\square\}_{\alpha\geq 0}$ and $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$ hold over $[0,+\infty)$:

Lemma 7.4.7 (analog of Lemma 7.4.3)**.** *For all $\alpha \geq 0$, we have $\tilde{V}^{\alpha/\tau}_M \subseteq P^\alpha_\square \subseteq \tilde{V}^{\tau\alpha}_M$.*

The proof of this result has the same flavor as the one of Lemma 7.4.3, but the details are slightly more technical due to the more elaborate definition of the filter $\tilde{t}$ and associated Voronoi filtration $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$.

Proof. Let x be a point of $\tilde{V}^{\alpha/\tau}_M \subseteq BB$, and let $v \in M$ be such that $x \in \tilde{U}_{\alpha/\tau}(v)$. If $v \in P$ with $\mathbf{d}_{M\setminus\{v\}}(v) > 2\alpha/\tau$, then $x \in \mathbf{ball}(v, \alpha/\tau)$ and thus $x \in P^{\alpha/\tau}_\square \subseteq P^\alpha_\square$. If $v \in P$ with $\mathbf{d}_{M\setminus\{v\}}(v) \leq 2\alpha/\tau$, then condition (iii) of Section 7.3.1 guarantees that $|x - v| \leq \alpha$ and thus $x \in P^\alpha_\square$. If $v \in S$, then the analysis is exactly the same as in the proof of Lemma 7.4.3. So, we have $\tilde{V}^{\alpha/\tau}_M \subseteq P^\alpha_\square$.

For the second inclusion, let x be a point of $P^\alpha_\square$. We want to show that $x \in \tilde{V}^{\tau\alpha}_M$. Let $v \in M$ be closest to x. If $v \in P$, then $x \in \mathbf{ball}(v, \alpha) \cap \mathrm{Vor}_\square(v)$, which is included in $\tilde{U}_\alpha(v)$ by definition (recall that we have $\tilde{t}(v) = 0$). As a result, $x \in \tilde{V}^{\tau\alpha}_M$. If $v \in S$, then the analysis is the same as in the proof of Lemma 7.4.3. So, we have $P^\alpha_\square \subseteq \tilde{V}^{\tau\alpha}_M$. □

We can now conclude the analysis in the same way as in Section 7.4.1. On the one hand, Lemma 7.4.6 and preceding discussion show that the filtrations $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$ and $\{\tilde{D}^\alpha_M\}_{\alpha\geq 0}$ have identical persistence diagrams. On the other hand, Lemma 7.4.4 implies that the filtrations $\{P^\alpha\}_{\alpha\geq 0}$ and $\{P^\alpha_\square\}_{\alpha\geq 0}$ have identical persistence diagrams. Finally, Lemma 7.4.7 shows a full (multiplicative) interleaving of $\{\tilde{V}^\alpha_M\}_{\alpha\geq 0}$ with $\{P^\alpha_\square\}_{\alpha\geq 0}$, which by Corollary 6.5.2 implies that their persistence diagrams are $\ln(\tau)$-close on the natural logarithmic scale. We thus obtain a stronger approximation guarantee than with the basic filter:

Theorem 7.4.8 (analog of Theorem 7.4.5)**.** *On the natural logarithmic scale, the persistence diagrams of $\{P^\alpha\}_{\alpha\geq 0}$ and $\{\tilde{D}^\alpha_M\}_{\alpha\geq 0}$ are $\ln\tau$-close in the bottleneck distance,*
i.e. $\mathbf{d}^{\ln}_{\mathrm{D}}\left(\{P^\alpha\}_{\alpha\geq 0},\ \{\tilde{D}^\alpha_M\}_{\alpha\geq 0}\right) \leq \ln\tau$.

7.5 Tighter Interleaving via Overmeshing

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a sizing function. As long as f is bounded from above by the Ruppert local feature size $\mathbf{f}_P$, NetMesh can return a mesh M such that the radius R_v of every Voronoi cell $\mathrm{Vor}(v)$ is at most $f(v)$. Given a parameter $\varepsilon > 0$, we will choose f be of the form $\frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P$. This means that for any $v \in M$ and any $x \in \mathrm{Vor}_\square(v)$, $|x - v| \leq \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(v)$. The standard mesh size

analysis implies that the size m of our output mesh M will be bounded as follows:

$$m = O\left(\int_{BB} \frac{1}{f(z)^d} dz\right) = O\left(\left(\frac{3(1+\varepsilon)}{\varepsilon}\right)^d \int_{BB} \frac{1}{\mathbf{f}_P(z)^d} dz\right). \tag{7.3}$$

In other words, our new sizing function f will only increase the mesh size by a factor of $\left(\frac{3(1+\varepsilon)}{\varepsilon}\right)^d$.

Modified α-mesh filtration. As before, we run the SVR algorithm on the input point set P, but this time using the sizing function f described above. Letting M denote the output superset of P, we modify the filter on Del_M in such a way that the Voronoi cells of mesh vertices that are significantly closer to a given point $p \in P$ than to the others appear only once p lies within $\alpha/2$ of its nearest neighbor in $P \setminus \{p\}$.

More precisely, for every point $x \in \mathbb{R}^d$ let n_x denote the point of P closest to x — if there are two or more such points, then choose either of them as n_x. We define the following function on the mesh vertices:

$$\forall v \in M,\ s'(v) = \max\left\{\mathbf{d}_P(v),\ \frac{1}{2}\mathbf{f}_P(n_v)\right\}.$$

Note that when v belongs to P, we have $n_v = v$ and $s'(v) = \frac{1}{2}\mathbf{f}_P(v)$. Also, if v is equidistant to two vertices $p, q \in P$, then $\mathbf{d}_P(v) \geq \frac{1}{2}\mathbf{f}_P(p)$ and $\mathbf{d}_P(v) \geq \frac{1}{2}\mathbf{f}_P(q)$, so the choice of which serves as n_v is irrelevant. Our new filter $t' : \mathrm{Del}_M \to \mathbb{R}$ is defined as follows:

- for each vertex v, let $t'(v) = 0$ if $v \in P$ and $t'(v) = s'(v)$ if $v \in M \setminus P$,
- $t'(\sigma) = \max_{i \in \{0,\cdots,k\}} s'(v_i)$ for each higher-dimensional simplex $\sigma = \{v_0, \cdots, v_k\}$.

The modified α-mesh filtration $\{D'^{\alpha}_M\}_{\alpha \geq}$ is defined as the sublevel filtration of t', so once again each space D'^{α}_M is a subcomplex of Del_M.

Approximation guarantee. Once again the analysis is done in terms of a dual filtration $\{V'^{\alpha}_M\}_{\alpha \geq 0}$, defined by $V'^{\alpha}_M = \bigcup_{v \in M} U'_\alpha(v)$, where

$$U'_\alpha(v) = \begin{cases} \emptyset & \text{if } v \in M \setminus P \text{ and } \alpha < s'(v), \\ \mathbf{ball}(v, \frac{\alpha}{1+\varepsilon}) & \text{if } v \in P \text{ and } \alpha < s'(v), \\ \mathrm{Vor}_\square(v) & \text{otherwise.} \end{cases}$$

Let $\mathcal{U}'_\alpha$ denote the collection of sets $\{U'_\alpha(v)\}_{v \in M}$. In contrast with Section 7.4, the sets $U'_\alpha(v) \in \mathcal{U}'_\alpha$ are not monotonically increasing with α, the problem being that $U'_\alpha(v) \subseteq U'_\beta(v)$ when $v \in P$ and $\alpha < s'(v) \leq \beta$. Nevertheless, for our choice of sizing field f the family $\{V'^{\alpha}_M\}_{\alpha \geq 0}$ is still a filtration (see Lemma 7.5.2 below). The proof of this fact relies on the following technical result:

Lemma 7.5.1. *For all $v \in P$, $\mathbf{ball}(v, \frac{s'(v)}{1+\varepsilon}) \subseteq \bigcup_{u \in M : |u-v| \leq s'(v)} \mathrm{Vor}_\square(u)$, where $\varepsilon \leq \frac{1}{2}$ is a user defined parameter that controls the sizing function for M.*

Proof. Assume for a contradiction that $\mathbf{ball}(v, \frac{s'(v)}{1+\varepsilon})$ intersects $\mathrm{Vor}_\square(u)$ for some $u \in M$ such that $|u - v| > s'(v) = \frac{1}{2}\mathbf{f}_P(v)$, and let x be a point in the intersection. Using the triangle inequality, the

quality of bound on the Voronoi cells, and the Lipschitz property of $\mathbf{f}_P$, we obtain:

$$|u-v| \le |v-x| + |x-u| \le \frac{s'(v)}{1+\varepsilon} + \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(u) \le \frac{1}{2(1+\varepsilon)}\mathbf{f}_P(v) + \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(u)$$
$$\le \mathbf{f}_P(v)\left(\frac{1}{2(1+\varepsilon)} + \frac{\varepsilon}{3(1+\varepsilon)}\right) + \frac{\varepsilon}{3(1+\varepsilon)}|u-v|,$$

which implies that $|u-v| \le \mathbf{f}_P(v)\left(\frac{\frac{1}{2(1+\varepsilon)}+\frac{\varepsilon}{3(1+\varepsilon)}}{1-\frac{\varepsilon}{3(1+\varepsilon)}}\right) = \frac{1}{2}\mathbf{f}_P(v)$, which contradicts our hypothesis. □

Lemma 7.5.2. *Given $\varepsilon \le \frac{1}{2}$, the family $\{{V'}_M^{\alpha}\}_{\alpha\ge 0}$ is a valid filtration.*

Proof. Let $v \in M$ and $\beta \ge \alpha \ge 0$. By definition, we have $U'_\alpha(v) \subseteq U'_\beta(v)$ unless $v \in P$ and $\alpha < s'(v) \le \beta$, which is the case we will now address. In this case, we have $U'_\alpha(v) = \mathbf{ball}(v, \frac{\alpha}{1+\varepsilon})$ and $U'_\beta(v) = \mathrm{Vor}_\square(v)$. Let S denote the set $M \cap \mathbf{ball}(v, s'(v))$. For every $u \in S$ we must have $v = n_u$, for otherwise the triangle inequality would imply that $\mathbf{f}_P(v) \le |v - n_u| < 2s'(v) = \mathbf{f}_P(v)$, a contradiction. As a result, $s'(u) = s'(v) \le \beta$, and thus $U'_\beta(u) = \mathrm{Vor}_\square(u)$. Then, Lemma 7.5.1 implies that $U'_\alpha(v) \subseteq \bigcup_{u\in S} \mathrm{Vor}_\square(u) = \bigcup_{u\in S} U'_\beta(u) \subseteq {V'}_M^{\beta}$. □

As in the previous sections, the filtration $\{{V'}_M^{\alpha}\}_{\alpha\ge 0}$ is interleaved multiplicatively with $\{P_\square^\alpha\}_{\alpha\ge 0}$:

Lemma 7.5.3. *Given $\varepsilon \le \frac{1}{2}$, for all $\alpha \ge 0$, ${V'}_M^{\alpha/(1+\varepsilon)} \subseteq P_\square^\alpha \subseteq {V'}_M^{\alpha(1+\varepsilon)}$.*

Proof. First we prove ${V'}_M^{\alpha/(1+\varepsilon)} \subseteq P_\square^\alpha$. Let x be a point in ${V'}_M^{\alpha/(1+\varepsilon)}$, and let $v \in M$ be such that $x \in U'_{\alpha/(1+\varepsilon)}(v)$. There are several cases to consider, depending on the value of α and on the location of v. In each case, the goal is to show that $\mathbf{d}_P(x) \le \alpha$.

Case 1: $\alpha/(1+\varepsilon) < s'(v)$**.** In this case we have $v \in P$ and $U'_{\alpha/(1+\varepsilon)}(v) = \mathbf{ball}(v, \alpha/(1+\varepsilon)^2)$, which gives $\mathbf{d}_P(x) \le \frac{\alpha}{(1+\varepsilon)^2} \le \alpha$.

Case 2: $\alpha/(1+\varepsilon) \ge s'(v)$**.** Since $\mathbf{f}_P$ is 1-Lipschitz, we have $\mathbf{f}_P(v) \le \mathbf{f}_P(n_v) + \mathbf{d}_P(v) \le 2s'(v) + s'(v) = 3s'(v)$. Hence,

$$\mathbf{d}_P(x) \le \mathbf{d}_P(v) + |x-v| \le s'(v) + f(v) \le s'(v) + \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(v) \le s'(v)(1+\varepsilon) \le \alpha.$$

Now we prove the other inclusion, namely $P_\square^\alpha \subseteq {V'}_M^{\alpha(1+\varepsilon)}$. Let x be a point in $P_\square^\alpha$, and let $v \in M$ be closest to x. Then, $x \in \mathrm{Vor}_\square(v)$, and we will show that either $x \in U'_{\alpha(1+\varepsilon)}(v)$ or $x \in U'_{\alpha(1+\varepsilon)}(n_v)$. If $\alpha(1+\varepsilon) \ge s'(v)$ then $U'_{\alpha(1+\varepsilon)}(v) = \mathrm{Vor}_\square(v)$, which contains x, so we may assume $\alpha(1+\varepsilon) < s'(v)$.

Case 1: $v \in P$**.** In this case we have $U'_{\alpha(1+\varepsilon)}(v) = \mathbf{ball}(v, \alpha)$, which contains x by hypothesis.

Case 2: $v \in M \setminus P$ **and** $s'(v) = \frac{1}{2}\mathbf{f}_P(n_v)$**.** As mentioned above, we may assume $\alpha < \frac{s'(v)}{1+\varepsilon}$ and thus $|x - n_x| \le \alpha < \frac{\mathbf{f}_P(n_v)}{2(1+\varepsilon)}$. The points x and v have a common nearest neighbor in P, because if

$n_x \neq n_v$ then we can derive the following contradiction:

$$\begin{aligned}\mathbf{f}_P(n_v) &\leq |n_v - n_x| \leq |n_v - v| + |v - x| + |x - n_x| \leq \mathbf{d}_P(v) + \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(v) + \alpha \\ &< \frac{1}{2}\mathbf{f}_P(n_v) + \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(n_v) + \frac{\varepsilon}{3(1+\varepsilon)}|v - n_v| + \frac{1}{2(1+\varepsilon)}\mathbf{f}_P(n_v) \\ &< \left(\frac{1}{2} + \frac{\varepsilon}{2(1+\varepsilon)} + \frac{1}{2(1+\varepsilon)}\right)\mathbf{f}_P(n_v) = \mathbf{f}_P(n_v).\end{aligned}$$

Now, since $\alpha(1+\varepsilon) < s'(v) = \frac{1}{2}\mathbf{f}_P(n_v) = \frac{1}{2}\mathbf{f}_P(n_x)$, we deduce that $U'_\alpha(n_x) = \mathbf{ball}(n_x, \alpha(1+\varepsilon))$, which contains x.

Case 3: $v \in M \setminus P$ **and** $s'(v) = \mathbf{d}_P(v)$**.** Again, we may assume $\alpha < \frac{s'(v)}{1+\varepsilon}$ and thus $|x - n_x| < \frac{s'(v)}{1+\varepsilon}$. However, we can derive the following contradiction proving this case impossible:

$$\begin{aligned}s'(v) = \mathbf{d}_P(v) &\leq |v - x| + |x - n_x| \\ &< \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(v) + \frac{1}{1+\varepsilon}s'(v) \leq \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{f}_P(n_v) + \frac{\varepsilon}{3(1+\varepsilon)}\mathbf{d}_P(v) + \frac{1}{1+\varepsilon}s'(v) \\ &< \frac{2\varepsilon}{3(1+\varepsilon)}s'(v) + \frac{\varepsilon}{3(1+\varepsilon)}s'(v) + \frac{1}{1+\varepsilon}s'(v) = s'(v).\end{aligned}$$

□

It follows from Lemma 7.5.3 and Corollary 6.5.2 that the persistence diagrams of $\{V'^{\alpha}_M\}_{\alpha\geq 0}$ and $\{P^{\alpha}_{\square}\}_{\alpha\geq 0}$ on the natural logarithmic scale are $\ln(1+\varepsilon)$-close (and therefore ε-close) to each other in the bottleneck distance. In addition, Lemma 7.4.4 tells us that $\{P^{\alpha}_{\square}\}_{\alpha\geq 0}$ and $\{P^{\alpha}\}_{\alpha\geq 0}$ have identical persistence diagrams. All that remains to be done now is relate $\{V'^{\alpha}_M\}_{\alpha\geq 0}$ to our simplicial filtration $\{D'^{\alpha}_M\}_{\alpha\geq 0}$. First, we prove that D'^{α}_M coincides with the nerve of the collection $\mathcal{U}'_\alpha$ (Lemma 7.5.5), which requires the following technical result:

Lemma 7.5.4. *For all $\alpha \geq 0$ and all $v \in P$, if $s'(v) > \alpha$ then $\mathbf{ball}(v, \frac{\alpha}{1+\varepsilon}) \cap U'_\alpha(u) = \emptyset$ for all other $u \in M$.*

Proof. Suppose for a contradiction that there exists some $u \in M$ such that $U'_\alpha(v) \cap U'_\alpha(u) \neq \emptyset$. If $U'_\alpha(u) = \mathbf{ball}(u, \frac{\alpha}{1+\varepsilon})$ then $u \in P$ and we get the following contradiction:

$$\mathbf{f}_P(v) \leq |u - v| \leq \frac{2\alpha}{1+\varepsilon} < 2\alpha < 2s'(v) = \mathbf{f}_P(v).$$

If $U'_\alpha(u) = \mathrm{Vor}_{\square}(u)$ then $s'(u) \leq \alpha < s'(v)$. By Lemma 7.5.1, if the Voronoi cell $\mathrm{Vor}_{\square}(u)$ intersects $U'_\alpha(v) = \mathbf{ball}(v, \frac{\alpha}{1+\varepsilon})$ then $|u - v| \leq s'(v)$. In this case, we get the following contradiction:

$$\begin{aligned}s'(v) = \frac{1}{2}\mathbf{f}_P(v) &\leq \frac{1}{2}|v - n_u| \leq \frac{1}{2}(|v - u| + |u - n_u|) \\ &\leq \frac{1}{2}(s'(v) + d_P(u)) \leq \frac{1}{2}(s'(v) + s'(u)) < s'(v).\end{aligned}$$

□

Lemma 7.5.5. *The complex D'^{α}_M coincides with the nerve of the cover $\mathcal{U}'_\alpha$ of V'^{α}_M.*

Proof. For an input vertex $p \in P$, we have $\{p\} \in D'^{\alpha}_M$ and $U'_\alpha(p) \neq \emptyset$ for any $\alpha \geq 0$. For a Steiner vertex $v \in M \setminus P$, $\{v\} \in D'^{\alpha}_M \Leftrightarrow s'(v) \geq \alpha \Leftrightarrow U'_\alpha(v) \neq \emptyset$.

Let $v \in P$ be a vertex such that $s'(v) < \alpha$ and thus $U'_\alpha(v) = \mathbf{ball}(v, \frac{\alpha}{1+\varepsilon})$. By Lemma 7.5.4, $\{v\}$ is the only simplex containing v in the nerve of
$U'\alpha$. particular $\mathbf{ball}(p, \frac{\alpha}{1+\varepsilon}) \cap \mathbf{ball}(q, \frac{\alpha}{1+\varepsilon}) = \emptyset$. Similarly, Lemma 7.5.1 and the fact that $\mathrm{Vor}_\square(v) \subset \mathbf{ball}(v, \frac{\alpha}{1+\varepsilon})$ imply that all the neighbors u of v in Del_M have $s'(u) < \alpha$ and therefore $U'_\alpha(u) = \emptyset$. It follows that $\{p\}$ is also the only simplex containing v in D'^{α}_M.

So, for any simplex $\sigma = \{v_0, \ldots, v_k\}$ with $k \geq 2$ that appears in D'^{α}_M or in the nerve of $\mathcal{U}'_\alpha$, we have $U'_\alpha(v_i) = \mathrm{Vor}_\square(v_i)$ for all $i = 1 \ldots k$. Hence, $\sigma \in D'^{\alpha}_M$ if and only if σ belongs to the nerve of $\mathcal{U}'_\alpha$. □

Lemma 7.5.5 suggests to use the Persistent Nerve Lemma 6.4.3 to conclude that $\{D'^{\alpha}_M\}_{\alpha\geq 0}$ and $\{V'^{\alpha}_M\}_{\alpha\geq 0}$ have identical persistence diagrams. Unfortunately, although the sets $U'_\alpha(v)$ are convex, they are not monotonically increasing with α, so they do not satisfy all the hypotheses of the Persistent Nerve Lemma. Consequently, we need to go through an intermediate filtration, $\{\mathcal{NU}''_\alpha\}_{\alpha\geq 0}$, where each space $\mathcal{NU}''_\alpha$ is defined as the nerve of the collection of sets $\mathcal{U}''_\alpha = \{U''_\alpha(v)\}_{v\in M}$ where $U''_\alpha(v) = U'_\alpha(v) \cap \mathrm{Vor}_\square(v)$. Let $V''^{\alpha}_M = \bigcup_{v\in M} U''_\alpha(v)$. Since the sets $U''_\alpha(v)$ are convex and monotonically increasing with α, the Persistent Nerve Lemma 6.4.3 implies that $\{\mathcal{NU}''_\alpha\}_{\alpha\geq 0}$ and $\{V''^{\alpha}_M\}_{\alpha\geq 0}$ have identical persistence diagrams. Now, the persistence diagram of $\{V''^{\alpha}_M\}_{\alpha\geq 0}$ is the same as the one of $\{V'^{\alpha}_M\}_{\alpha\geq 0}$, by the following result:

Lemma 7.5.6. *For all $\alpha \geq 0$, the canonical inclusion $V''^{\alpha}_M \hookrightarrow V'^{\alpha}_M$ is a homotopy equivalence.*

Proof. We will exhibit a deformation retraction of V'^{α}_M onto V''^{α}_M. On each connected component of V'^{α}_M separately. By Lemma 7.5.4, every vertex $v \in P$ with $s'(v) > \alpha$ has the property that $U'_\alpha(v)$ is disjoint from all other sets $U'_\alpha(u)$ and thus forms a separate connected component. On this component the deformation retraction is easily defined using the metric projection onto the convex set $U''_\alpha(v)$, as in Lemma 7.4.4. All other connected components of V'^{α}_M can be expressed as unions of $U'_\alpha(u)$'s, each of which is equal to $\mathrm{Vor}_\square(u)$. For these components the identity map is a trivial deformation retraction. □

In addition, the persistence diagram of $\{\mathcal{NU}''_\alpha\}_{\alpha\geq 0}$ is the same as the one of $\{D'^{\alpha}_M\}_{\alpha\geq 0}$:

Lemma 7.5.7. *For all $\alpha \geq 0$, $\mathcal{NU}''_\alpha = D'^{\alpha}_M$.*

Proof. Observe that $U''_\alpha(v) \subseteq U'_\alpha(v)$ for all points $v \in M$, so $\mathcal{NU}''_\alpha$ is naturally included in the nerve of $\mathcal{U}'_\alpha$, which by Lemma 7.5.5 coincides with D'^{α}_M. For the other inclusion, we observe that $U'_\alpha(v) \subseteq U''_\alpha(v)$ unless $v \in P$ and $\alpha < s'(v)$. However, by Lemma 7.5.4, such vertices v only appear in 0-simplices of D'^{α}_M. These 0-simplices also appear in $\mathcal{NU}''_\alpha$ since $\mathrm{Vor}_\square(v) \cap \mathbf{ball}(v, \frac{\alpha}{1+\varepsilon}) \neq \emptyset$, so indeed, $D'^{\alpha}_M \subseteq \mathcal{NU}''_\alpha$. □

It follows from Lemmas 7.5.6 and 7.5.7 that $\{D'^{\alpha}_M\}_{\alpha\geq 0}$ and $\{V'^{\alpha}_M\}_{\alpha\geq 0}$ have identical persistence diagrams. This concludes our analysis and gives our main theoretical result:

Theorem 7.5.8. *Given any user-defined parameter $\varepsilon \in (0, \frac{1}{2}]$ controlling the sizing function for M, the persistence diagrams of $\{P^{\alpha}\}_{\alpha\geq 0}$ and $\{D'^{\alpha}_M\}_{\alpha\geq 0}$ on the natural logarithmic scale are ε-close in the bottleneck-distance, i.e. $\mathbf{d}^{\ln}_{\mathrm{D}}(\{P^{\alpha}\}_{\alpha\geq 0},\ \{D'^{\alpha}_M\}_{\alpha\geq 0}) \leq \varepsilon$.*

7.6 Recursively Well-Paced Subsets

7.6.1 Partitioning the input by scale

If α is significantly bigger than the distance between two points p and q in P, then intuitively, the difference between the offsets P^{α} and $(P \setminus \{q\})^{\alpha}$ should also be small. In this section, we make this intuition precise and show that the persistence diagram of the union of the clipped offsets of the sets $\{P_i\}$ is close to the persistence diagram of P^{α}.

Recall that the **bounding ball** of P_i is $B_i = \mathbf{ball}(p_i, 2r_i)$. The **clipped offset** of P_i is defined to be the set $P^{\alpha}_{i\circ} = P^{\alpha}_i \cap B_i$. The union of these clipped offsets is defined to be $P^{\alpha}_{*\circ} = \bigcup_i P^{\alpha}_{i\circ}$.

Let the **extended bounding ball** of P_i be $B_{i+} = \mathbf{ball}(p_i, \frac{1-\theta}{2\theta} r_i)$. We will also define the **extended clipped offsets** $P^{\alpha}_{i+} = P^{\alpha}_i \cap B_{i+}$. As with the clipped offsets, we define $P^{\alpha}_{*+} = \bigcup_i P^{\alpha}_{i+}$. We can also consider the extended clipped offsets of the entire set P, $P^{\alpha}_{+} = P^{\alpha} \cap B_{1+}$. The extended offsets are useful because they are clipped less aggressively than the regular clipped offsets, but, as we will see, they have the same topology. This usefulness is demonstrated in the following Lemma.

Lemma 7.6.1. *For all $\alpha \geq 0$, $P^{\alpha/(1+3\theta)}_{*+} \subseteq P^{\alpha}_{+} \subseteq P^{\alpha(1+3\theta)}_{*+}$.*

Proof. For the first inclusion, let x be a point in $P^{\alpha/(1+3\theta)}_{*+}$. So, for some i, $x \in P^{\alpha/(1+3\theta)}_i \cap B_{i+}$. Since $P_i \subseteq P$, we have that $P^{\alpha/(1+3\theta)} \subseteq P^{\alpha}$. Moreover, $B_{i+} \subseteq B_{1+}$ so $x \in P^{\alpha} \cap B_{1+} = P^{\alpha}_{+}$.

To prove the second inclusion, let x be any point in P^{α}_{+}. We will show that $x \in P^{\alpha(1+3\theta)}_{*+}$. Let i be the maximum such that $x \in B_{i+}$. Let n_x be the nearest neighbor to x in P. Since $x \in P^{\alpha}$, we know that

$$|x - n_x| \leq \alpha. \tag{7.4}$$

If $n_x \in P_i$ then x is contained in $P^{\alpha}_i \cap B_{i+} = P^{\alpha}_{i+}$, which is itself contained in $P^{\alpha(1+3\theta)}_{*+}$. So, we may assume that $n_x \notin P_i$.

Choose j to be the largest such that $n_x \in C_j$ and $p_j \in P_i$. Since $n_x \in C_j$, the definition of r_j implies that

$$|n_x - p_j| \leq r_j. \tag{7.5}$$

By our choice of i, we know that $x \notin B_{j+}$ and thus

$$r_j < \frac{2\theta}{1-\theta}|x - p_j|. \tag{7.6}$$

We can show that x and p_j are close using the triangle inequality as well as the inequalities from equations (7.4), (7.5), and (7.6) as follows.

$$|x - p_j| \leq |x - n_x| + |n_x - p_j| < \alpha + r_j \leq \alpha + \frac{2\theta}{1-\theta}|x - p_j| \leq \alpha(1 + 3\theta).$$

The last inequality uses the fact that $\theta \leq \frac{1}{9}$. □

Lemma 7.6.2. *The log-scale persistence diagrams of $\{P^\alpha_{*\circ}\}$ and $\{P^\alpha\}$ are 3θ-close in bottleneck distance.*

Proof. By Lemma 7.4.4, $P^\alpha \hookrightarrow P^\alpha_+$ is a homotopy equivalence, and so $\{P^\alpha\}$ and $\{P^\alpha_+\}$ have identical persistence diagrams. The preceding lemma and Corollary 6.5.2 imply that $\mathbf{d}^{\ln}_{\mathrm{D}}(\{P^\alpha_+\}, \{P^\alpha_{*+}\}) \leq \log(1+3\theta) \leq 3\theta$. So, it will suffice to show that $\{P^\alpha_{*+}\}$ and$\{P^\alpha_{*\circ}\}$ have identical persistence diagrams, which we prove by showing that the inclusion map $P^\alpha_{*\circ} \hookrightarrow P^\alpha_{*+}$ is a homotopy equivalence.

The difference between $P^\alpha_{*\circ}$ and P^α_{*+} is in the size of the balls used around each point set P_i. Let A_i be the annulus between these two balls, i.e. $A_i = B_{i+} \setminus B_i$. The choice of $\frac{1-\theta}{\theta} r_i$ as the outer radius was precisely chosen so that the A_i's are pairwise disjoint (this is the only place we use the assumption that the r_is are distinct). We use the metric projection $\pi_{P^\alpha_{*\circ}} : P^\alpha_{*+} \to P^\alpha_{*\circ}$ piecewise on the A_i. For any x, if $\pi_{P^\alpha_{*\circ}}(x) \neq x$ then $x \in A_i$ for some i and $\pi_{P^\alpha_{*\circ}}(x) = \frac{(x-p_i)\max\{2r_i,\alpha\}}{|x-p_i|}$. From the definition, $\pi_{P^\alpha_{*\circ}}$ is also clearly 1-Lipschitz with the annuli. We can now follow exactly the same arguments as in Lemma 7.4.4 to construct the homotopy equivalence and complete the proof. □

7.6.2 Approximation Guarantee

We now prove that the α-mesh filtration, $\{D^\alpha_{M_*}\}_{\alpha\geq 0}$, is a good approximation to the offsets filtration for computing persistence diagrams. First we will show that the persistence diagram of the α-Voronoi filtration, $\{V^\alpha_{M_*}\}_{\alpha\geq 0}$, is close to that of the offsets of P. This will be a straightforward application of the results we have proved thus far. Second, we will show that $\{V^\alpha_{M_*}\}_{\alpha\geq 0}$ and $\{D^\alpha_{M_*}\}_{\alpha\geq 0}$ have identical persistence diagrams.

Lemma 7.6.3. *For all $\alpha \geq 0$ and each mesh M_i, $V^{\alpha/(1+\varepsilon)}_{M_i} \subseteq P^\alpha_{i\circ} \subseteq V^{\alpha(1+\varepsilon)}_{M_i}$.*

Proof. The interleaving can be shown by the exact same arguments as in Lemma 7.5.3. □

Lemma 7.6.4. $\mathbf{d}^{\ln}_{\mathrm{D}}(\{V^\alpha_{M_*}\}_{\alpha\geq 0}, \{P^\alpha\}_{\alpha\geq 0}) \leq 3\theta + \varepsilon$.

Proof. We can apply the interleaving of Lemma 7.6.3 to each M_i to find that $V^{\alpha/(1+\varepsilon)}_{M_*} \subseteq P^\alpha_\circ \subseteq V^{\alpha(1+\varepsilon)}_{M_*}$. So, by Corollary 6.5.2, we have $\mathbf{d}^{\ln}_{\mathrm{D}}(\{V^\alpha_{M_*}\}_{\alpha\geq 0}, \{P^\alpha_{*\circ}\}_{\alpha\geq 0}) \leq \log(1+\varepsilon) \leq \varepsilon$. Applying Lemma 7.6.2 and the triangle inequality completes the proof. □

Let $D^\alpha_i = \bigcup_{j=1}^i D^\alpha_{M_j}$. So, in particular, $D^\alpha_k = D^\alpha_{M_*}$ and $D^\alpha_i \subseteq D^\alpha_{M_*}$ for all $i = 1 \ldots k$.

Lemma 7.6.5. *If $i \leq j$ and $\alpha \geq r_i$, then $D^\alpha_i \hookrightarrow D^\alpha_j$ is a homotopy equivalence.*

Proof. We proceed by induction on j. The base case of $j = i = 1$ is trivial because then $D_i^\alpha = D_j^\alpha = D_1^\alpha$. By induction, assume that the statement holds for $j' < j$. It will suffice to show that $D_{j-1}^\alpha \hookrightarrow D_j^\alpha$ is a homotopy equivalence and the desired homotopy equivalence can be constructed by composition.

Since $\alpha \geq r_i > r_j$, it must be that $D_{M_j}^\alpha = \mathrm{Del}_{M_j}$ and is thus homeomorphic to a ball and contractible. The complex $D_{M_j}^\alpha$ intersects D_{j-1}^α only at the vertex p_j. By contracting $D_{M_j}^\alpha$ to p_j we get that $D_{j-1} \hookrightarrow D_j$ is a homotopy equivalence. Thus, by the inductive hypothesis and the composition of homotopies, we get that $D_i \hookrightarrow D_j$ is a homotopy equivalence as desired.

□

When α is in the range $[r_{i+1}(1+\varepsilon), r_i(1+\varepsilon)]$, the complex D_i^α is an embedded simplicial complex. More importantly, the following Lemma shows that it is the nerve of a cover of $V_{M_*}^\alpha$ defined by the sets $\overline{U}_\alpha(v) = U_\alpha^{j_v}(v)$, where j_v is defined to be the maximum integer less than or equal to i such that $v \in M_{j_v}$.

Lemma 7.6.6. *For any $i \in [0, k]$ and all α such that $r_{i+1} \leq \frac{\alpha}{1+\varepsilon} \leq r_i$, D_i^α coincides with the nerve of $\{\overline{U}_\alpha(v)\}_{v \in M_{[i]}}$.*

Proof. First observe that $\overline{U}_\alpha \neq \emptyset$ if and only if $\{v\} \in D_i^\alpha$ so it will suffice to show that the two complexes coincide for simplices of dimension ≥ 1.

First, we show that any simplex in D_i^α corresponds to a nonempty intersection of $\overline{U}_\alpha(v)$'s. Suppose we have $\sigma \in D_i^\alpha$. Then we have $\sigma \in D_{M_j}^\alpha$ for some $j \leq i$ and so by the definition of $D_{M_j}^\alpha$, it is the case that $\bigcap_{v \in \sigma} U_\alpha^j(v) \neq \emptyset$. It will suffice to $j_v = j$ for all $v \in \sigma$. Clearly, $v \in M_j$ for each vertex $v \in \sigma$, so if $j_v \neq j$ then $j_v > j$. Since v is in some 1-simplex in $D_{M_j}^\alpha$, $\alpha \geq d_{P_j} \geq \frac{1-\theta}{\theta} r_{j_v} > r_j(1+\varepsilon)$. By our hypothesis, this implies that $j_v > i$. However, this is impossible by the definition of j_v.

Now, we will show that if $\bigcap_{v \in \sigma \subset M_{[i]}} \overline{U}(v) \neq \emptyset$ then there is a corresponding simplex in D_i^α. It will suffice to show that if for any two vertices u, v, $\overline{U}(u) \cap \overline{U}(v) \neq \emptyset$ then $j_u = j_v$, because this will imply that $\overline{U}(v) = U_\alpha^j(v)$ for a fixed value of j and thus σ corresponds to a simplex in $D_{M_j}^\alpha \subseteq D_i^\alpha$. Suppose for contradiction that $j_v < j_u$ for some such pair. Then, by our hypothesis, $\frac{\alpha}{1+\varepsilon} \leq r_{j_u}$. Choose $x \in \overline{U}(u) \cap \overline{U}(v)$. The set $\overline{U}(u)$ is contained in the bounding ball B_{j_u}, so $|x - p_{j_u}| \leq r_{j_u}$. Let n_v be the nearest point in P to v as before. Because $\overline{U}(v)$ is nonempty, $|v - n_v| \leq \alpha$ By the interleaving, we know that $\overline{U}(v) \subseteq P^{\alpha(1+\varepsilon)}$ and so $|x - n_v| \leq \alpha(1+\varepsilon) \leq r_{j_u}(1+\varepsilon)^2$. Recalling that $\varepsilon \leq \frac{1}{2}$, we can combine these with the triangle inequality, to obtain that $|p_{j_u} - n_v| \leq r_{j_u}(1 + (1+\varepsilon)^2) \leq \frac{13}{4} r_{j_u}$. On the other hand, the well-paced construction guarantees that $|p_{j_u} - n_v| \geq \frac{1-\theta}{2\theta} r_{j_u}$. Recalling that $\theta < \frac{1}{9}$, this implies that $4r_{j_u} = \frac{1-\frac{1}{9}}{2\frac{1}{9}} r_{j_u} \leq |p_{j_u} - n_v| \leq \frac{13}{4} r_{j_u}$, a contradiction.

□

Lemma 7.6.7. *For any $i \in [0, k]$ and all α, β such that $r_{i+1} \leq \alpha \leq \beta \leq r_i$, the following diagram*

commutes at the homology level.

$$\begin{array}{ccc} V^{\alpha}_{M_*} & \hookrightarrow & V^{\beta}_{M_*} \\ \downarrow & & \downarrow \\ D^{\alpha}_{i} & \hookrightarrow & D^{\beta}_{i} \\ \downarrow & & \downarrow \\ D^{\alpha}_{M_*} & \hookrightarrow & D^{\beta}_{M_*} \end{array}$$

Proof. The top half commutes at the homology level by Lemma 7.6.6 and the Persistent Nerve Lemma. The bottom half commutes because all of the maps are inclusions and thus it induces a commutative diagram at the homology level as well. □

Lemma 7.6.8. $\mathrm{D}\{D^{\alpha}_{M_*}\}_{\alpha\geq 0} = \mathrm{D}\{V^{\alpha}_{M_*}\}_{\alpha\geq 0}$

Proof. We need to show the existence of homomorphisms to make the following diagram commute where the horizontal maps are induced by inclusion.

$$\begin{array}{ccc} H_k(V^{\alpha}_{M_*}) & \to & H_k(V^{\beta}_{M_*}) \\ \downarrow & & \downarrow \\ H_k(D^{\alpha}_{M_*}) & \to & H_k(D^{\beta}_{M_*}) \end{array}$$

This follows from Lemma 7.6.7 by observing that the inclusion map $V^{\alpha}_{M_*} \hookrightarrow V^{\beta}_{M_*}$ can be decomposed into $V^{\alpha}_{M_*} \hookrightarrow V^{r_i}_{M_*} \hookrightarrow \cdots \hookrightarrow V^{r_j}_{M_*} \hookrightarrow V^{\beta}_{M_*}$, and similarly for $D^{\alpha}_{M_*} \hookrightarrow D^{\beta}_{M_*}$. □

We can now state the main Theorem, which follows directly from the preceding Lemmas and the triangle inequality.

Theorem 7.6.9. *The persistence diagrams of* $\{P^{\alpha}\}_{\alpha\geq 0}$ *and* $\{D^{\alpha}_{M_*}\}_{\alpha\geq 0}$ *on the natural logarithmic scale are* $(3\theta+\varepsilon)$*-close in the bottleneck distance, i.e.* $\mathbf{d}^{\ln}_{\mathrm{D}}(\{D^{\alpha}_{M_*}\}_{\alpha\geq 0}, \{P^{\alpha}\}_{\alpha\geq 0}) \leq 3\theta+\varepsilon$.

7.7 Experiments

As a proof of concept, we applied the approach of Section 7.4 to 2,000 points sampled on the 4-dimensional Clifford torus, as described in Figure 7.2. We modified a pre-existing SVR implementation [AHMP07] to run in 4D and compute the filtration of Section 7.4.2. We used the Plex library [ZZZ] to compute the persistence diagram. To the 2,000 input data points, SVR added approximately 71,000 Steiner points including a bounding box and achieved an aspect ratio bound of $\tau = 3.08$ (a value chosen for technical reasons related to the bounding box). In total, the mesh contained about 2 million pentahedra, 12 million simplices overall. It took approximately 1 hour to compute the mesh and filtration, and another 7 hours to compute the persistence diagram.

Figure 7.2 (right) displays the persistence diagram thus obtained as a **persistence barcode** [CZCG04]: Homological features are sorted first by their dimension, then by their start time, and drawn as an interval. The interval with an arrow head with arrow heads extends to infinity. The qualitative interpretation of the barcode is straightforward: scanning through the

scales from smallest to largest, we see the point cloud, the helicoidal curve, the Clifford torus, the 3-sphere of radius $\sqrt{2}$, and finally the ambient space $\mathbb{R}^4$, represented simply as a space with trivial reduced homology groups. Note that the topological noise appearing in the 2-dimensional barcode between -0.2 and 0 is made of many short intervals of length less than 0.05. The 3-sphere structure is of particular interest because it had never been observed before, being too far from the beginning of the filtration for Rips or Čech filtration techniques to capture it.

Quantitatively, the curve appears at time $\tau\alpha = -1.73$, which corresponds roughly to half the distance between consecutive points along the curve. The second 1-cycle of the torus appears around $\tau\alpha = -1.2$, which is only slightly sooner than the time ($\tau\alpha = -1.16$) at which consecutive periods of the curve start being connected in the offsets filtration. The 2-cycle of the torus appears soon afterwards, since the square $[0, 2\pi]^2$ gets filled in rapidly once consecutive periods of the curve start being connected. The isolines $u = \mathrm{C}^{\mathrm{t}}$ and $v = \mathrm{C}^{\mathrm{t}}$ are mapped to unit circles in $\mathbb{R}^4$, so both the 1-cycles and the 2-cycle should disappear at $\tau\alpha = \tau 1 = 0$ in the barcode, which is close to being the case. Among the points that lie farthest away from the Clifford torus on the 3-sphere, we have $(\sqrt{2}, 0, 0, 0)$, whose their distance to the torus is $\sqrt{4 - 2\sqrt{2}} \approx 1.08$, so the 3-sphere should appear at $\tau\alpha = \tau 1.08 \approx 0.08$ in the barcode, which it does approximately. At the end of the barcode the approximation quality worsens a bit: since the 3-sphere has radius $\sqrt{2}$, the 3-cycle should disappear at $\tau\alpha = \tau\sqrt{2} \approx 0.35$, but in reality it does so sooner, around $\tau\alpha = 0.18$. Nevertheless, the absolute error is still within $\tau 1.18$, meaning that our result is as good as if a multiplicative 1.18-interleaving had been obtained, whereas the aspect ratio bound τ used by the SVR algorithm was 3.08 (a value chosen for technical reasons relating to SVR's handling of the bounding box). So, it appears from this analysis that the quality of approximation provided by our method can be significantly better in practice than expected from the theory.

Comparison. The 4-skeleton of the Rips filtration of P reaches an equivalent size (2 million pentahedra) as early as $\tau\alpha = -0.75$, which makes it difficult with this budget to detect the torus, and impossible to detect the 3-sphere. Increasing the limit to a mere $\tau\alpha = -0.5$ already raises the size of the Rips filtration to more than 10 million simplices. The Clifford torus is not a worst case for the α-complex filtration. However, as mentioned, the α-complex is susceptible to pathological behavior on some other very reasonable inputs.

Engineering issues. Our implementation is very preliminary and would benefit from substantial engineering. In particular, the SVR implementation on which we based our filtering software adds points to a bounding box to avoid dealing with Steiner points near the boundary of space. The number of points on the bounding box is negligible in two and three dimensions, but outnumbered our input for the four-dimensional example. The bounding box also limited the quality we could practically achieve. In addition, since we did not have access to efficient staged predicates and constructions in 4D, we used exact rational arithmetic, which in 3D slows SVR down by a factor of worse than 20. Despite this, meshing was not the bottleneck compared to the persistence computation using Plex.

7.8 Discussion

Steiner point choice. Since all our filtrations are derived from the mesh Del_M, their sizes (and therefore the complexity of the whole approach) depend heavily on the size of M. Some work has been done in two and three dimensions to optimize point placement (e.g. [Üng09]), reducing the mesh size for any requested quality, or alternately in practice allowing better quality than is theoretically achievable (i.e. allowing meshing to $\tau < 2$). Furthermore, there is a huge industry in mesh smoothing, which in practice improves the quality of a mesh as a post-processing step. Reductions in the number of Steiner points are particularly important as the dimension increases, whereas improving the quality improves the approximation.

Higher dimension. A major limitation of our approach resides in the fact that it is tied to the ambient space $\mathbb{R}^d$, which is fine in small to moderate dimensions but not in high dimensions. One possibility for improvement would be to refine the approach and its analysis, so as to make its complexity depend on the dimensionality of the topological features the user is interested in. For instance, in scenarios where the data are high-dimensional but are known to lie on or close to low-dimensional geometric structures of low dimensions, it would be interesting to devise a mechanism that allows the user to capture the low-dimensional topological features at all scales, at a cost that does not depend exponentially on the ambient dimension. Some work has been done in this direction [CO08], mainly using Rips or witness complex filtrations, but it remains preliminary for the moment. It would be interesting to see if meshing techniques could help in this context.

7.9 Technical Lemmas

Let K be a closed convex set in $\mathbb{R}^d$, and let π_K denote the metric projection onto K,

$$\pi_K(x) = \operatorname*{argmin}_{y \in K} |x - y|.$$

Lemma 7.9.1. *For any compact convex set $K \subseteq \mathbb{R}^d$, the projection π_K is well-defined and 1-Lipschitz over all of $\mathbb{R}^d$.*

Proof. First, we show that π_K is well-defined. Let x be any point in $\mathbb{R}^d$. Suppose for contradiction that there were two points $a, b \in K$ such that

$$|x - a| = |x - b| = \mathbf{d}_K(x).$$

Let $c = \frac{a+b}{2}$ be the point midway between a and b. Since K is convex, $c \in K$. However, $|x - c| < |x - a|$, a contradiction.

We now show that π_K is 1-Lipschitz, i.e. $|\pi_K(x) - \pi_K(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}^d$. Let x and y be any two points in $\mathbb{R}^d$ and let $x' = \pi_K(x)$ and $y' = \pi_K(y)$ be their projections onto K. We may assume $x' \neq y'$ for otherwise $|x' - y'| \leq |x - y|$ is trivial. Let L be the line through x' and y'. The

points $x'' = \pi_L(x)$ and $y'' = \pi_L(y)$ are the projections of x and y onto L. The distance between x'' and y'' may be bounded as follows.

$$|x'' - y''| = \left(\frac{x' - y'}{|x' - y'|}\right)^T (x - y) \leq |x - y|.$$

The point x'' cannot lie in the line segment $\overline{x'y'}$, because then x'' would be closer to x than x'. Similarly, $y'' \notin \overline{x'y'}$. Also, x'' and y'' lie on opposite sides of $\overline{x'y'}$ because $x' \neq y'$. So, $|x' - y'| \leq |x'' - y''| \leq |x - y|$ as desired.

□

Lemma (7.4.4)**.** *Let B be a compact, convex set. If $Q \subset B$ then for all $\alpha \geq 0$, the canonical inclusion $Q^\alpha \cap B \hookrightarrow Q^\alpha$ is a homotopy equivalence.*

Proof. Let π_B denote the metric projection onto B, $\pi_B(x) = \operatorname{argmin}_{y \in B}|x - y|$. Lemma 7.9.1 ensures that π_B is well-defined and 1-Lipschitz. Let x be a point of Q^α, and let $x' = \pi_B(x)$.

We will show that the line segment $\overline{x, x'}$ is included in Q^α. Let $p \in Q$ be such that $|x - p| \leq \alpha$. Since B contains Q, we have $\pi_B(p) = p$, and therefore $|p - x'| \leq |p - x|$ since π_B is 1-Lipschitz. It follows that both x and x' belong to $\mathbf{ball}(p, \alpha)$. Since this ball is convex, it contains the whole line segment $\overline{x, y}$.

We will now construct a homotopy explicitly. Define $F : [0, 1] \times Q^\alpha \to \mathbb{R}^d$ as $F(t, x) = (1 - t)x + t\pi_B(x)$. Since π_B is 1-Lipschitz, F is continuous. In addition, the above discussion shows that $F(t, Q^\alpha) \subseteq Q^\alpha$ for all $t \in [0, 1]$. Also, since $Q^\alpha \cap B \subseteq B$, the restriction of π_B to $Q^\alpha \cap B$ is the identity, as is the restriction of F. Finally, for all $x \in Q^\alpha$ we have $F(1, x) = \pi_B(x) \in Q^\alpha \cap B$. Hence, F is a deformation retraction of Q^α onto $Q^\alpha \cap B$, which implies that the canonical inclusion $Q^\alpha \cap B \hookrightarrow Q^\alpha$ is a homotopy equivalence. □

Chapter 8

Geometric Persistent Homology for General Distance Functions

8.1 Overview

A distance function can induce a shape from a point cloud P in $\mathbb{R}^d$. For example, the α-offsets of P is the set of points in $\mathbb{R}^d$ of distance at most α from P. Much is known about the geometry and topology of offsets as well as how to compute them. As estimators of the shape, the offsets fail in the presence of background noise and outliers.

To combat this problem, alternative distance functions can be used such as the kth nearest neighbor distance. Rather than choosing k explicitly, we show how to build a multifiltered complex that captures both the scale α and the smoothing parameter k. The persistence vineyard of this multifiltration gives a picture of what topological features may be present in the data as well as their robustness to change in both α and k.

To avoid the blowup in size of the underlying complex, one can also compute an approximation. The goal is to produce a filtered complex of small size whose persistence diagram is provably close to that produced by the true (k, α)-offsets. For this, we give a general approach to geometric approximation of sublevel filtrations in constant dimensional Euclidean space. We show how to $(1 + \varepsilon)$-approximate the persistence diagram of the sublevel filtration of any Lipschitz function $f \geq \frac{d_2}{c}$ for constant $c > 0$, where $d_2(x)$ is the distance from x to its 2nd nearest neighbor in P. Moreover, if we have an ordered sequence of such functions, we can compute a single complex that can be filtered for all of them. This is done by extending the mesh-based filtration algorithm introduced in our previous work with Hudson, Miller, and Oudot [HMOS10]. As special cases, this works for the kth nearest neighbor distance as well as the k-distances introduced by Chazal, Cohen-Steiner, and Merigot [CCSM10].

8.2 Offsets and De-noising

The beauty of persistent homology for topological data analysis is that it obviates the need to choose an explicit scale at which to view the data. Instead, the algebraic structure of the problem allows one to compute for all possible scales simultaneously. Not only does one skip the problem of tuning parameters, but also the output shows explicitly which features are robust to perturbations of the parameters. Unfortunately, de-noising the data often leads to new parameters to choose. We show how to replace the Euclidean distance with a family of distance functions to de-noise the data as part of the persistence computation. The result is a vineyard of persistence diagrams that captures not only what topological features are present but also how robust they are to changes in the de-noising parameter. We give exact constructions for the kth nearest neighbor distance and a general approximation method for a broad class of distance-like functions.

In practice, geometric persistent homology has three phases: one statistical, one geometric, and the third, topological [CIdSZ08]. First, the data is filtered for noise. Second, the geometry of the points drives the construction of a filtered simplicial complex. Third, the persistent homology of the filtered complex is computed. Usually, the emphasis is placed on the latter two phases with the first treated as a necessary evil. And it *is* necessary; even a small number of outliers can generate spurious persistent features that foil existing methods.

Traditionally, the process of de-noising the data introduces a new set of parameters that must be tuned explicitly, one for the scale at which to define density and one to be the threshold between signal and noise. So, although no explicit scale is chosen to compute the persistent homology, one *is* chosen to de-noise the data. The problem is both aesthetic and practical. Not only would it be more elegant to do all three phases without tuning any parameters, it would also be more useful, as it has been observed that de-noising parameters can be difficult to choose [KC10].

We solve this problem by folding the noise removal phase into the persistence computation. The result is a family of simplicial complexes filtered in two dimensions, capturing both the scale of the data and the threshold for noise. This setting is closely related to density-based clustering in which one may choose a smoothing parameter for a density estimator as well as a noise threshold.

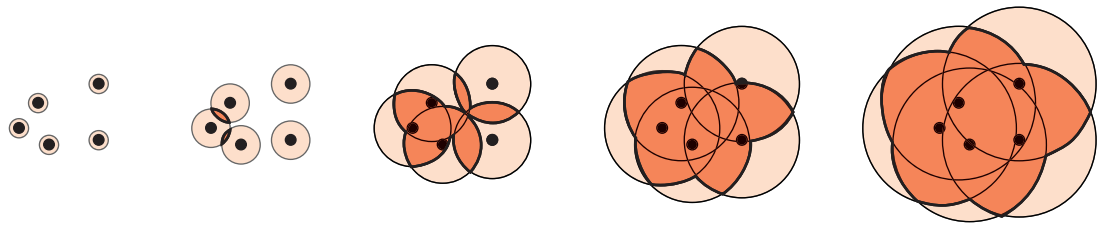

Figure 8.1: The α-offsets overlaid with the $(2, \alpha)$-offsets.

In the first part of this chapter, we replace the usual distance to the input set P, with the kth nearest neighbor distance, d_k. This replaces the α-offsets, $P^\alpha = \bigcup_{p \in P} \mathbf{ball}(p, \alpha)$ with the (k, α)-**offsets**, $P_k^\alpha = d_k^{-1}(-\infty, \alpha]$ as our estimate of the shape at scale α (see Figure 8.1). This is closely related to a common approach to de-noising data for topological data analysis points are treated

as noise if the distance to their kth nearest neighbor is at least some threshold α (see [NSW08b] and [CIdSZ08] for two notable examples). We show how to filter by α rather than fixing it in advance.

We can now also vary k, which corresponds intuitively to the number of points that must participate locally for a feature to be considered signal rather than noise. This is related to another standard de-noising heuristic in which connected components corresponding to individual input points are ignored, regardless of their persistence. The kth nearest neighbor distance takes that intuition and encodes it in the distance function in a way that also applies to higher order features generated by more than a single point (see Figure 8.2).

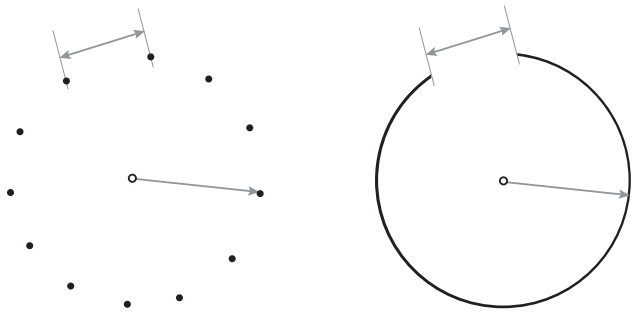

Figure 8.2: Two inputs yielding persistent 1-cycles with the same birth and death times. The cycle on the left uses relatively few points while the cycle on the right uses many. Our approach allows us to tell the difference between these two cases.

Contribution 1 We present **barycentric multifiltrations**, a natural way to filter a simplicial complex by a second parameter. We then show that the barycentric mutifiltration of the Čech filtration captures exactly the persistent homology of the (k, α)-offsets. This is then related to the Rips complex.

Next, we turn our attention to approximation methods for a wider class of "distance-like" functions. We extend the method of filtering triangulations generated by Voronoi refinement meshing. These have the advantage that under mild sampling assumptions on the input, they maintain size that is linear in the number of points.

Contribution 2 We prove approximation guarantees for mesh-based filtrations of more general smooth functions. This allows us to construct a multifiltered complex that approximates the (k, α)-offsets that is the same size as the corresponding mesh filtration for the α-offsets. Perhaps surprisingly, for $k > 1$, both the filtration and its analysis are simpler than for the usual distance function, d_1. Thus these results are both simpler to understand and implement while also being more generally applicable.

8.3 Related Work

They give a simple update procedure to connect two filtrations on the same complex.

Independent of persistent homology, the problem of homology inference from discrete input has also been the subject of much interest. Niyogi, Smale, and Weinberger showed that for sufficiently dense samples on smooth manifolds, the offsets are in fact, homotopy equivalent to the underlying manifold [NSW08a]. Chazal and Lieutier extended these results to the case of non-uniform and noisy samples [CL08, CCSL09]. Both of these method are based on offsets.

Chazal et al. have looked at more exotic distance functions to do geometric inference in the presence of noise, achieving very strong guarantees [CCSM10]. The issue of topological de-noising has also been addressed by Kloke and Carlsson [KC10]. They present several examples where the direct application of k-nearest neighbor distance de-noising fails because of the difficulty of picking the relevant parameters. This is precisely the problem we propose to attack by folding the selection of these parameters into the persistence algorithm itself, rather than treating the de-noising step as an independent pre-process.

Filtering by the sublevels of d_k differs from the use of the kth nearest neighbor graph is used in pattern recognition (see [Mar04] for a survey) or in nonlinear dimensionality reduction (see [BN01, RS00, dST02, TdS00] among others). In that context, the assumption is that the points are sampled nearly uniformly from some unknown distribution so that areas of low density indicate that the intrinsic metric has been stretched by the embedding. With the kth nearest neighbor distance function, we are assuming only that the sample is "dense enough" in the input space. This means that we do not infer information about the underlying metric structure, but we also require less stringent sampling conditions in order to discover the relevant structure.

8.4 Background

Distance Functions and Offsets. Let $P \subset \mathbb{R}^d$ be a set of distinct points. We can generalize the offsets of P by changing the distance function. Define the **kth nearest neighbor distance**, d_k, to be the distance to k points of P. That is,

$$d_k(x) = \min_{S \in \binom{P}{k}} \max_{p \in S} |x - p|.$$

The sublevels of the kth nearest neighbor distance are the (k, α)-offsets, denoted P_k^α:

$$P_k^\alpha = d_k^{-1}(-\infty, \alpha].$$

Equivalently, the **(k, α)-offsets** are the points contained in at least k balls of radius α centered at points in P (see Figure 8.1). Note that d_1 is exactly d_P and d_2 is the Ruppert feature size $\mathbf{f}_P$.

k-distances and distance-like functions. As shown by Merigot, a large range of "distance-like" functions can be used to infer topology [Mér10]. Chazal et al. [CCSM10] and Kloke and Carlsson [KC10] independently observed that the gradient of the distance function seems to be especially useful in topological de-noising. This was made concrete in the notion of a distance to a measure, which for point sets yields the k-distance:

$$\mathrm{D_k}(x)^2 = \frac{1}{k} \sum_{p \in kNN(x)} |p - x|^2,$$

where $kNN(x)$ denotes the k nearest neighbors of x among the points of P [CCSM10]. Recently, Guibas et al. demonstrated a method for computing k-distances using a small number of witness points [GMM11]. In that paper, they give a coarser approximation than ours, but achieve smaller complexity wih respect to the ambient dimension. As a corollary to our main theorem on mesh-based filtrations, we will show that the persistent homology of the sublevel filtration of $\mathrm{D_k}$ can be approximated over all scales. Moreover, the complex can be filtered by k as well.

8.5 Barycentric Multifiltration

As always, let $P \subset \mathbb{R}^d$ be a set of n points. Let $S = \{S_p \subset \mathbb{R}^d : p \in P\}$ be a family of closed sets indexed by the points of P. For any subset u of P, let T_u be the intersection of the corresponding sets:

$$T_u = \bigcap_{p \in u} S_p.$$

Let $\mathcal{J}$ be the nerve S realized with vertices in P:

$$\mathcal{J} = \{u \subset P : T_u \neq \emptyset\}.$$

Let $\tilde{\mathcal{J}}$ be the barycentric decomposition of $\mathcal{J}$:

$$\tilde{\mathcal{J}} = \big\{\{u_1, u_2, \ldots\} \subset \mathcal{J} : u_1 \subset u_2 \subset \cdots\big\}.$$

Finally, we define $\tilde{\mathcal{J}}_k$ as the subcomplex of $\tilde{\mathcal{J}}$ induced on the vertices of cardinality at least k:

$$\tilde{\mathcal{J}}_k = \{\sigma \in \tilde{\mathcal{J}} : \text{for all } v \in \sigma, |v| \geq k\}.$$

Note that in the special case where $S_p = \overline{\mathbf{ball}}(p, \alpha)$, $\tilde{\mathcal{J}}_k = \tilde{\mathcal{C}}_k^\alpha$.

We define the **k-nerve** of S to be the nerve of k-wise intersections among the sets of S:

$$\mathcal{N}_k = \left\{ U \subseteq \binom{P}{k} : \bigcap_{u \in U} T_u \neq \emptyset \right\}.$$

The barycentric decomposition of the k-nerve is denoted $\tilde{\mathcal{N}}_k$:

$$\tilde{\mathcal{N}}_k = \big\{\{u_1, u_2, \ldots\} \subset \mathcal{N}_k : u_1 \subset u_2 \subset \cdots\big\}.$$

The complex $\tilde{\mathcal{N}}_k$ contains a lot of redundant information. Let V be the vertex set of $\tilde{\mathcal{N}}_k$. We will define a map $\pi : V \to V$ that "projects" out a lot of this redundant information.

$$\pi(v) = \binom{\bigcup_{u \in v} u}{k}.$$

We will use this map to compare some of the complexes defined above in Lemmas 8.5.1 and 8.5.3.

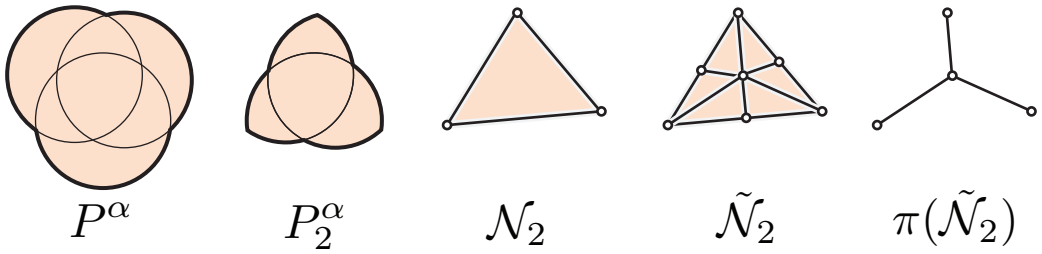

Figure 8.3: The construction of the 2-nerve, its barycentric decomposition, and its image under π.

Lemma 8.5.1. *For any finite collection of closed sets S, the complexes $\tilde{\mathcal{J}}_k$ and $\pi(\tilde{\mathcal{N}}_k)$ associated with S are isomorphic.*

Proof. We define a map ϕ from the vertex set of $\tilde{\mathcal{J}}_k$ to the vertex set of $\pi(\tilde{\mathcal{N}}_k)$ as $\phi(u) = \binom{u}{k}$. The inverse of this map is $\phi^{-1}(v) = \bigcup_{v_i \in v} v_i$. So, ϕ takes subsets $u \subset P$ of size at least k such that $T_u \neq \emptyset$ to the family of k-element subsets of u. It is easy to check that ϕ is a bijection. It will suffice to prove that σ is a simplex of $\tilde{\mathcal{J}}_k$ if and only if $\phi(\sigma)$ is a simplex of $\pi(\tilde{\mathcal{N}}_k)$. Let $\sigma = (u_0, \ldots, u_j) \in \tilde{\mathcal{J}}_k$ be any simplex. By the definition of $\tilde{\mathcal{J}}_k$, $u_0 \subset \cdots \subset u_j$. For any pair of vertices u_a and u_b, $u_a \subset u_b$ if and only if $\phi(u_a) \subset \phi(u_b)$. So $\sigma \in \tilde{\mathcal{J}}_k$ if and only if $\phi(u_0) \subset \cdots \subset \phi(u_j)$, which holds if and only if $\phi(\sigma) \in \pi(\tilde{\mathcal{N}}_k)$. □

Before proceeding to prove the homotopy equivalence between $\pi(\tilde{\mathcal{N}}_k)$ and $\tilde{\mathcal{N}}_k$, we give state a minor technical lemma. It gives a sufficient combinatorial condition for the existence of a deformation retraction between two simplicial complexes. It's proof is a simple exercise in homotopy theory and may be found in Section 8.8.

Lemma 8.5.2. *Let $\mathcal{K}_1 \subset \mathcal{K}_2$ be simplicial complexes with vertex sets V_1 and V_2. Let $\pi : V_2 \to V_1$ be a map that restricts to the identity on V_1. If $\pi(\sigma) \in \mathcal{K}_1$ and $\sigma \cup \pi(\sigma) \in \mathcal{K}_2$ for all $\sigma \in \mathcal{K}_2$ then $\mathcal{K}_1$ is a deformation retract of $\mathcal{K}_2$.*

Lemma 8.5.3. *The complex $\pi(\tilde{\mathcal{N}}_k)$ is a deformation retract of $\tilde{\mathcal{N}}_k$.*

Proof. Let V be the set of vertices of $\tilde{\mathcal{N}}_k$ and let $s = \max_{v \in V} |v|$. Let for $i = 0 \dots s$, define

$$V_i = \{v \in V : |v| \leq i\} \cup \{\pi(v) : v \in V, |v| > i\}$$

Let A_i be the subcomplex of $\tilde{\mathcal{N}}_k$ induced on V_i. So

$$\pi(\tilde{\mathcal{N}}_k) = A_0 \subset \cdots \subset A_s = \tilde{\mathcal{N}}_k.$$

It will suffice to show that the inclusion $A_i \hookrightarrow A_{i+1}$ is a homotopy equivalence for all i. Lemma 8.5.2 gives a recipe for constructing such a homotopy equivalence from a map between the vertex sets. Let $\pi_i : V_i \to V_{i-1}$ be defined as

$$\pi_i(v) = \begin{cases} v & \text{if } |v| < i \\ \pi(v) & \text{if } |v| \geq i \end{cases}$$

Lemma 8.5.2 will give the desired deformation retraction as long as the following conditions are met.

(1) π_i restricts to the identity on V_{i-1},
(2) $\pi_i(\sigma) \in A_{i-1}$ for all $\sigma \in A_i$, and
(3) $(\sigma \cup \pi_i(\sigma)) \in A_i$ for all $\sigma \in A_i$.

Item (1) is obvious from the definitions. To prove (2) and (3), fix a simplex $\sigma = (v_0, \dots, v_t) \in A_i$ and let $\sigma' = \sigma \cup \pi_i(\sigma)$. If $\sigma = \sigma'$ then we are done, so we may assume that for some vertex $v_j \in \sigma$, $\pi(v_j) \notin \sigma$. Recall that the simplices of $\tilde{\mathcal{N}}_k$ (and also A_i) are strictly nested sequences of vertices. So, there is at most one such vertex v_j, namely the one with cardinality i. We may therefore express σ' as $\sigma \cup \{\pi(v_j)\}$. Since $\sigma \in A_i \subset \tilde{\mathcal{N}}_k$,

$$v_0 \subset \cdots \subset v_t.$$

Observe that $v \subseteq \pi(v)$ for all $v \in V$ and moreover that $u \subset v$ if and only if $\pi(u) \subset \pi(v)$. So, it follows that

$$v_0 \subset \cdots \subset v_j \subset \pi(v_j) \subset \pi(v_{j+1}) = v_{j+1} \subset \cdots \subset v_t.$$

This is a strictly nested sequence of the vertices of σ' so $\sigma' \in A_i$, proving (3). Moreover, $\pi(\sigma) = \sigma' \setminus \{v_j\}$ so $\pi(\sigma) \in \tilde{\mathcal{N}}_k$ as well. Since $\pi(\sigma) \subset V_{i-1}$, we conclude that $\pi(\sigma) \in A_{i-1}$, proving (2). □

The preceding lemmas are combinatorial statements and make no assumption about the topology of the sets of S. The next theorem introduces such a topological hypothesis, requiring the intersections to be well-behaved. It may be viewed as a generalization of the Nerve Theorem.

Theorem 8.5.4. *Let $S = \{S_1, \dots, S_n\}$ be a family of sets. For any subset $u \subset S$, let $T_u = \bigcap_{S_i \in u} S_i$.*

Let $T = \{T_u : u \in \binom{P}{k}\}$. If T_u is empty or contractible for each $u \subset S$, then

$$\bigcup_{T_u \in T} T_u \simeq \tilde{\mathcal{J}}_k.$$

Proof. We first observe that $\mathcal{N}_k$ is the nerve of T. So the Nerve Theorem implies that $\bigcup_{T_u \in T} T_u \simeq \mathcal{N}_k$. A complex and its barycentric decomposition are homeomorphic so $\mathcal{N}_k \simeq \tilde{\mathcal{N}}_k$. Combining these facts with Lemmas 8.5.1 and 8.5.3 yields

$$\bigcup_{T_u \in T} T_u \simeq \mathcal{N}_k \simeq \tilde{\mathcal{N}}_k \simeq \pi(\tilde{\mathcal{N}}_k) \simeq \tilde{\mathcal{J}}_k.$$

□

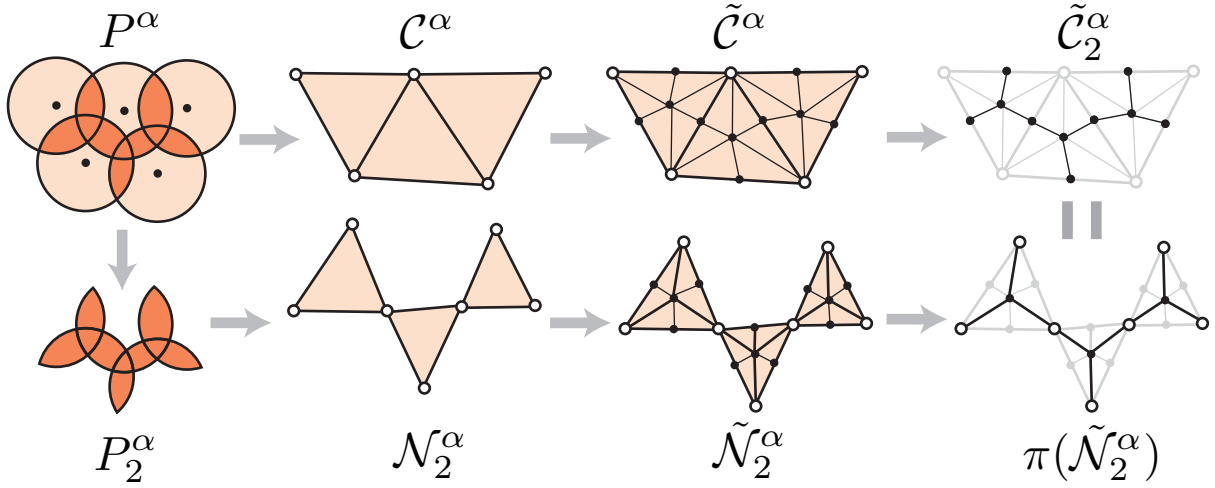

Figure 8.4: We transform the collection of balls in two different ways to get equivalent complexes, $\tilde{\mathcal{C}}_k^\alpha$ (top) and $\pi(\tilde{\mathcal{N}}_k^\alpha)$ (bottom) for $k = 2$.

The barycentric Čech complex. The vertices of the barycentric decomposition $\tilde{\mathcal{K}}$ of a complex $\mathcal{K}$ are simplices of $\mathcal{K}$; they are sets. This leads to a natural filtration on $\tilde{\mathcal{K}}$ defined to be $\{\tilde{\mathcal{K}}_k\}_k$, where $\tilde{\mathcal{K}}_k$ is the subcomplex induced on the vertices of cardinality at least k (the filter parameter here goes down rather than up but this is not a problem). If we have a filtered complex $\{\mathcal{K}^\alpha\}_\alpha$, then we can apply this method to form the **barycentric multifiltration**, $\{\tilde{\mathcal{K}}_k^\alpha\}_{k,\alpha}$. Let $\tilde{\mathcal{C}}^\alpha$ be the barycentric decomposition of the Čech complex at scale α, and let $\{\tilde{\mathcal{C}}_k^\alpha\}_{k,\alpha}$ be its barycentric multifiltration. The following theorem establishes the equivalence of the persistence diagrams of $\{\tilde{\mathcal{C}}_k^\alpha\}$ and $\{P_k^\alpha\}$.

Theorem 8.5.5. *For any fixed k, the persistence diagram of the barycentric Čech filtration, $\{\tilde{\mathcal{C}}_k^\alpha\}$, is identical to the persistence diagram of the (k, α)-offsets, $\{P_k^\alpha\}$.*

Proof. It suffices to demonstrate a family isomorphisms $\{\tilde{\mathcal{C}}_k^\alpha \to P_k^\alpha\}_{\alpha \geq 0}$ that commute at the homology level with the canonical inclusions. We let $S_\alpha = \{S_p = \overline{\mathbf{ball}}(p, \alpha) : p \in P\}$ be our collection of closed sets. For this choice of S, $\tilde{\mathcal{J}}_k = \tilde{\mathcal{C}}_k^\alpha$. Let $\mathcal{N}_k^\alpha$ be the k-nerve of S_α. We will show that the following diagram commutes at the homology level, where the vertical maps are canonical inclusions and the horizontal maps induce isomorphisms at the homology level.

$$\begin{array}{ccccccccc}
\tilde{\mathcal{C}}_k^\alpha & \to & \pi(\tilde{\mathcal{N}}_k^\alpha) & \hookrightarrow & \tilde{\mathcal{N}}_k^\alpha & \to & \mathcal{N}_k^\alpha & \to & P_k^\alpha \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\tilde{\mathcal{C}}_k^\beta & \to & \pi(\tilde{\mathcal{N}}_k^\beta) & \hookrightarrow & \tilde{\mathcal{N}}_k^\beta & \to & \mathcal{N}_k^\beta & \to & P_k^\beta
\end{array}$$

The first box commutes because Lemma 8.5.1 implies that $\tilde{\mathcal{C}}_k^\alpha$ and $\pi(\tilde{\mathcal{N}}_k^\alpha)$ are isomorphic simplicial complexes. The second box commutes because Lemma 8.5.3 implies that the inclusion is a homotopy equivalence. The third box commutes because $\tilde{\mathcal{N}}_k^\alpha$ and $\mathcal{N}_k^\alpha$ are homeomorphic for all α and the homeomorphism $\tilde{\mathcal{N}}_k^\beta \to \mathcal{N}_k^\beta$ restricts to the homeomorphism $\tilde{\mathcal{N}}_k^\alpha \to \mathcal{N}_k^\alpha$ on $\tilde{\mathcal{N}}_k^\alpha$. The fourth box commutes by the Persistent Nerve Lemma and the observation that $\tilde{\mathcal{N}}_k^\alpha$ is the nerve of a good cover of P_k^α. □

The barycentric Rips complex Let $\tilde{\mathcal{R}}^\alpha$ denote the barycentric decomposition of the Rips complex, $\mathcal{R}^\alpha$. The following theorem shows that the barycentric Rips complex can be used to approximate the persistence diagram of the (k, α)-offsets in the same way that the Rips complex is used for α-offsets.

Theorem 8.5.6. *For any fixed k, the persistence diagram of the barycentric Rips filtration, $\{\tilde{\mathcal{R}}_k^\alpha\}$, is a $\sqrt{2}$-approximation to the persistence diagram of the (k, α)-offsets $\{P_k^\alpha\}$ when the underlying space is Euclidean, and is a 2-approximation for general metrics.*

Proof. Let c be the interleaving constant equal to $\sqrt{2}$ in Euclidean space and 2 for general metrics. It is known that the Rips and Čech complexes are interleaved, $\mathcal{C}^\alpha \subseteq \mathcal{R}^\alpha \subseteq \mathcal{C}^{c\alpha}$ (see [dSG07a] for a proof along with a slightly tighter dimension-dependent bound). This extends directly to an interleaving between the barycentric Rips and barycentric Čech complexes. That is, $\tilde{\mathcal{C}}_k^\alpha \subseteq \tilde{\mathcal{R}}_k^\alpha \subseteq \tilde{\mathcal{C}}_k^{c\alpha}$. This interleaving, Theorem 6.5.3, and Theorem 8.5.5 imply the desired approximation guarantee. □

8.6 Filtrations on Meshes

A clear drawback of the barycentric Čech multifiltration is that its size blows up quite rapidly. In this section, we show how to use meshes to achieve a good approximation to the persistence diagram. To accomplish this, we extend the work of the previous chapter to give a general condition for which the mesh-based approach gives good approximatin guarantees. In the end, we will see that not only can we get approximations for kth nearest neighbor distances but also for any sufficiently large Lipschitz function.

Let M be a τ-well-spaced superset of the inputs set P. That is, Vor_M is a τ-quality mesh.

For any function $f : \mathbb{R}^d \to \mathbb{R}$, we can construct three natural filtrations based on f, Vor_M, and Del_M:

1. The **sublevel filtration**: $\{F^\alpha\}$, where
$$F^\alpha = f^{-1}(-\infty, \alpha].$$
2. The **Voronoi filtration**: $\{\mathrm{V}^\alpha\}$, where
$$\mathrm{V}^\alpha = \bigcup_{\substack{v \in M \\ f(v) \leq \alpha}} \mathrm{Vor}(v).$$
3. The **Delaunay Filtration**: $\{\mathcal{D}^\alpha\}$, where
$$\mathcal{D}^\alpha = \{\sigma \in \mathrm{Del}_M : \forall v \in \sigma, f(v) \leq \alpha\}.$$

Observe that each $\mathcal{D}^\alpha$ is the nerve of the family of convex sets $\{\mathrm{Vor}(v) : v \in M, f(v) \leq \alpha\}$ and thus the Delaunay filtration is homotopy equivalent to the Voronoi filtration. We will follow the pattern that algorithms operate on the Delaunay filtration and proofs work with the Voronoi filtration.

Given a function f that is sufficiently large and sufficiently smooth, the Delaunay filtration on a τ-quality Delaunay triangulation gives a constant factor approximation to the sublevel filtration of f. Below, we discuss specific sufficient conditions for f and prove the corresponding approximation guarantee in Theorem 8.6.3.

We start with a lemma that compares a function to an approximation that is constant on Voronoi cells. A good approximation yields an interleaving of the corresponding filtrations, $\{F^\alpha\}$ and $\{\mathrm{V}^\alpha\}$.

Lemma 8.6.1. *If for all $v \in M$ and all $x \in \mathrm{Vor}(v)$, $\frac{1}{c} f(x) \leq f(v) \leq c f(x)$ for some constant $c \geq 1$, then $\mathrm{V}^{\alpha/c} \subseteq F^\alpha \subseteq \mathrm{V}^{c\alpha}$, for all $\alpha \geq 0$.*

Proof. First we prove that $\mathrm{V}^{\alpha/c} \subseteq F^\alpha$. If x is a point in $\mathrm{V}^{\alpha/c}$ then $f(v) \leq \alpha/c$. It follows that $f(x) \leq \alpha$, and so $x \in F^\alpha$. Next, we prove that $F^\alpha \subseteq \mathrm{V}^{c\alpha}$. If x is in F^α then $f(x) \leq \alpha$ and thus $f(v) \leq c\alpha$. It follows that $\mathrm{Vor}(v) \subset \mathrm{V}^{c\alpha}$, and so $x \in \mathrm{V}^{c\alpha}$. □

8.6.1 An ε-Refined Mesh

We will show that a mesh can be used to approximate the sublevel filtration of any sufficiently large Lipschitz function f. We can compute the mesh M without regard to f except for its Lipschitz constant and the constant $c > 0$ such that $f \geq \frac{d_2}{c}$. Moreover, the filtration will be $\mathcal{D}^\alpha$, so filtering the mesh just requires that we evaluate the function at the vertices.

Definition. *A mesh M is ε-**refined** if for every mesh vertex v in M, the outer radius R_v is at most $\varepsilon d_2(v)$.*

If the sample set P is well-paced the NetMesh algorithm augmented with the over-refinement process defined in Section 3.7 can compute this linear size mesh $O(n \log n)$ time. The constants will depend on ε.

Lemma 8.6.2. *If M is ε-refined and f is a t-Lipschitz function with $f \geq \frac{d_2}{c}$ for some constant $c > 0$, then*

$$\frac{1}{1+\varepsilon_0} f(x) \leq f(v) \leq (1+\varepsilon_0) f(x).$$

for all $v \in M$ and $x \in \mathrm{Vor}(v)$, where $\varepsilon_0 = \frac{ct\varepsilon}{1-ct\varepsilon}$.

Proof. Let $v \in M$ and $x \in \mathrm{Vor}(v)$ be chosen arbitrarily. Then we may bound $f(x)$ as follows.

$$\begin{aligned}
f(x) &\leq f(v) + t|v-x| && [\text{f is } t\text{-Lipschitz}] \\
&\leq f(v) + t\varepsilon d_2(v) && [R_v \leq \varepsilon d_2(v)] \\
&\leq c(1+t\varepsilon) f(v) && [d_2 \leq cf] \\
&< (1+\varepsilon_0) f(v). && [\varepsilon_0 > ct\varepsilon]
\end{aligned}$$

Similarly, we can bound $f(v)$:

$$\begin{aligned}
f(v) &\leq f(x) + t|v-x| && [f \text{ is } t\text{-Lipschitz}] \\
&\leq f(x) + t\varepsilon d_2(v) && [R_v \leq \varepsilon d_2(v)] \\
&\leq f(x) + ct\varepsilon f(v) && [d_2 \leq cf] \\
&\leq \frac{1}{1-ct\varepsilon} f(x) && [\text{Collect terms}] \\
&= (1+\varepsilon_0) f(x) && \left[1+\varepsilon_0 = \frac{1}{1-ct\varepsilon}\right]
\end{aligned}$$

□

The preceding lemmas and Theorem 6.5.3 imply the following theorem.

Theorem 8.6.3. *If M is an ε-refined mesh, and $f \geq \frac{d_2}{c}$ is t-Lipschitz with $c > 0$, then the persistence diagram of the Delaunay (or equivalently, the Voronoi) filtration on f and M is a $\frac{1}{1-ct\varepsilon}$-approximation to the persistence diagram of the sublevels filtration of f.*

As before, the result for general functions applies easily to the class kth nearest neighbor distances to yield the following theorem.

Corollary 8.6.4. *If M is ε-refined and $k \geq 2$, then the persistence diagram of the Delaunay (or equivalently, the Voronoi) filtration on d_k and M is a $\frac{1}{1-\varepsilon}$-approximation to the persistence diagram of the sublevels filtration of f. Similarly, the persistence diagram of the Delaunay (or equivalently,*

the Voronoi) filtration on $\mathrm{D}_{\mathrm{k}k}$ *and* M *is a* $\frac{1}{1-\sqrt{2}\varepsilon}$*-approximation to the persistence diagram of the sublevels filtration of* f*.*

Proof. First observe that d_k and D_{k} are both 1-Lipschitz. Next observe that for $k \geq 2$, $d_k \geq d_2$ and $\mathrm{D}_{\mathrm{k}k} \geq \frac{d_2}{\sqrt{2}}$. The latter follows because

$$\mathrm{D}_{\mathrm{k}k}^2 \geq \mathrm{D}_{\mathrm{k}2}^2 = \frac{1}{2}d_1 + \frac{1}{2}d_2 \geq \frac{1}{2}d_2.$$

So the corollary follows from Theorem 8.6.3. □

8.7 Conclusions and directions for future work

We have presented a set of filtrations to extend offset methods in persistent homology to kth nearest neighbor distance functions. We showed how the barycentric decomposition of the Čech and Rips complexes directly induces a filter parameterized k. We also showed how mesh-based methods can be extended to provide strong guarantees for a wide class of functions. In particular, a single ε-refined mesh can be filtered to give a $1+\varepsilon$-approximation to the (k,α)-offsets for any $k \geq 2$, yielding an efficient method for computing persistent homology in the presence of noise.

Another direction for future work is to extend this work to other statistical density estimators or other distance functions. In both cases, there are multiple relevant variables that may be incorporated into multifiltrations.

8.8 Technical Lemmas

Lemma (8.5.2)**.** *Let* $\mathcal{K}_1 \subset \mathcal{K}_2$ *be simplicial complexes with vertex sets* V_1 *and* V_2*. Let* $\pi : V_2 \to V_1$ *be a map that restricts to the identity on* V_1*. If* $\pi(\sigma) \in \mathcal{K}_1$ *and* $\sigma \cup \pi(\sigma) \in \mathcal{K}_2$ *for all* $\sigma \in \mathcal{K}_2$ *then* $\mathcal{K}_1$ *is a deformation retract of* $\mathcal{K}_2$*.*

Proof. The complexes $\mathcal{K}_1$ and $\mathcal{K}_2$ have a natural embedding as subcomplexes of the standard simplex in $\mathbb{R}^D$ where $D = |V_2|$. We use underline to denote vertices, simplices, and complexes in this embedding, i.e. $\underline{v}$, $\underline{\sigma}$, and $\underline{\mathcal{K}_1}$. It will suffice to demonstrate a deformation retraction from $\underline{\mathcal{K}_2}$ to $\underline{\mathcal{K}_1}$.

For any $x \in \underline{\mathcal{K}_2}$, let σ_x denote the unique minimal simplex of $\mathcal{K}_2$ such that $x \in \mathrm{conv}(\underline{\sigma_x})$. The coefficients $\{\lambda_v : v \in \sigma_x\}$ that realize this convex combination are the **barycentric coordinates** of x:

$$x = \sum_{v \in \sigma_x} \lambda_v \underline{v}, \qquad \forall v \in \sigma_x : \lambda_v > 0, \qquad \sum_{v \in \sigma_x} \lambda_v = 1.$$

Using the barycentric coordinates, we define the map $\pi_* : \underline{\mathcal{K}_2} \to \underline{\mathcal{K}_1}$ by

$$\pi_*(x) = \sum_{v \in \sigma_x} \lambda_v \underline{\pi(v)}.$$

This map is realized in $\mathbb{R}^D$ by a linear projection, and therefore is continuous. We also observe that the image of π_* does lie in $\underline{\mathcal{K}_1}$ as asserted because for any $x \in \underline{\mathcal{K}_2}$, $\pi_*(x) \in \text{conv}(\underline{\pi(\sigma_x)})$ and $\pi(\sigma_x) \in \mathcal{K}_1$ by the hypothesis of the lemma.

We now define the linear homotopy F as

$$F(t, x) = (1 - t)x + t\pi_*(x).$$

Observe that for all $x \in \underline{\mathcal{K}_2}$, $F(0, x) = x$ and $F(1, x) = \pi_*(x) \in \underline{\mathcal{K}_1}$. The homotopy is continuous because π_* is continuous (in fact, it can be realized by an affine map in $\mathbb{R}^{D+1}$).

To show that F is a deformation retraction, it suffices to show that $F(t, x) \in \underline{\mathcal{K}_2}$ for all $t \in [0, 1]$. Since $x \in \text{conv}(\underline{\sigma_x})$ and $\pi_*(x) \in \text{conv}(\underline{\pi(x)})$,

$$F(t, x) \in \text{conv}(\underline{\sigma_x} \cup \underline{\pi(\sigma_x)}) = \text{conv}(\underline{\sigma_x \cup \pi(\sigma_x)}).$$

The hypothesis of the lemma says that $\sigma_x \cup \pi(\sigma_x) \in \mathcal{K}_2$ so $F(t, x) \in \underline{\mathcal{K}_2}$ as desired. □

Bibliography

[AdG08] Boris Aronov, Mark de Berg, and Chris Gray. Ray shooting and intersection searching amidst fat convex polyhedra in 3-space. *Computational Geometry: Theory and Applications*, 41:68–76, 2008. 5.2

[AHMP07] Umut A. Acar, Benoît Hudson, Gary L. Miller, and Todd Phillips. SVR: Practical engineering of a fast 3D meshing algorithm. In *Proc. 16th International Meshing Roundtable*, pages 45–62, 2007. 4.15, 7.7

[BA76] Ivo Babuška and A. K. Aziz. On the Angle Condition in the Finite Element Method. *SIAM Journal on Numerical Analysis*, 13(2):214–226, April 1976. 1.1

[Bar02] Alexander Barvinok. *A Course in Convexity*. American Mathematical Society, 2002. 2

[BDadHS05] Marcin Bienkowski, Valentina Damerow, Friedhelm Meyer auf der Heide, and Christian Sohler. Average case complexity of Voronoi diagrams of n sites from the unit cube. In *EWCG: European Workshop Comput. Geom.*, pages 167–170, 2005. 5.2

[BEG94] Marshall Wayne Bern, David Eppstein, and John Russell Gilbert. Provably good mesh generation. *J. Computer & Systems Sciences*, 48(3):384–409, June 1994. Special issue for 31st FOCS. 4.2

[BET99] M. Bern, D. Eppstein, and S.-H. Teng. Parallel construction of quadtrees and quality triangulations. *International Journal of Computational Geometry and Applications (IJCGA)*, 9(6):517–532, 1999. 4.2

[BN01] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in Neural Information Processing Systems 14*, pages 585–591, 2001. 8.3

[Car09] Gunnar Carlsson. Topology and data. *Bull. Amer. Math. Soc.*, 46:255–308, 2009. 1.2

[CBK09] Moo K. Chung, Peter Bubenik, and Peter T. Kim. Persistence diagrams of cortical surface data. In *Information Processing in Medical Imaging*, volume 5636, pages 386–397, 2009. 1.2

[CCSG+09] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J. Guibas, and Steve Y. Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the 25th ACM Symposium on Computational Geometry*, pages 237–246, 2009. 6.5, 6.5.1, 6.5

[CCSL09] Frédéric Chazal, David Cohen-Steiner, and André Lieutier. A sampling theory for compact sets in euclidean space. *Discrete & Computational Geometry*, 41:461–479, 2009. 8.3

[CCSM10] Frédéric Chazal, David Cohen-Steiner, and Quentin Mérigot. Geometric Inference for Measures based on Distance Functions. Research Report RR-6930, INRIA, http://hal.inria.fr/inria-00383685/PDF/RR-6930v2.pdf, 2010. 8.1, 8.3, 8.4

[CDE+00] Siu-Wing Cheng, Tamal K. Dey, Herbert Edelsbrunner, Michael A. Facello, and Shang-Hua Teng. Sliver exudation. *JACM: Journal of the ACM*, 47, 2000. 1.1, 3.5

[CDR05] Siu-Wing Cheng, Tamal K. Dey, and Tathagata Ray. Weighted Delaunay refinement for polyhedra with small angles. In *IMR*, pages 325–342, 2005. 3.5

[CDRR05] Siu-Wing Cheng, Tamal K. Dey, Edgar A. Ramos, and Tathagata Ray. Quality meshing of polyhedra with small angles. *International Journal of Computational Geometry and Applications (IJCGA)*, 15(4):421–461, 2005. 3.5

[CdS10] Gunnar Carlsson and Vin de Silva. Zigzag persistence. *Foundations of Computational Mathematics*, 10(4):367–405, 2010. 6.3

[CdSM09] Gunnar Carlsson, Vin de Silva, and Dmitriy Morozov. Zigzag persistent homology and real-valued functions. In *Proceedings of the 25th ACM Symposium on Computational Geometry*, pages 247–256, 2009. 6.3

[Cha00] Bernard Chazelle. *The Discrepancy Method*. Cambridge University Press, 2000. 4.7.1, 4.7.5

[Che89] L. Paul Chew. Guaranteed-quality triangular meshes. Technical Report TR-89-983, Department of Computer Science, Cornell University, 1989. 3.2

[CIdSZ08] Gunnar Carlsson, Tigran Ishkhanov, Vin de Silva, and Afra Zomorodian. On the local behavior of spaces of natural images. *International Journal of Computer Vision*, 76(1):1–12, 2008. 1.2, 8.2, 8.2

[CL07] Frédéric Chazal and André Lieutier. Stability and computation of topological invariants of solids in R^n. *Discrete & Computational Geometry*, 37(4):601–617, 2007. 6.3

[CL08] Frédéric Chazal and André Lieutier. Smooth manifold reconstruction from noisy and non-uniform approximation with guarantees. *Computational Geometry: Theory and Applications*, 40:156–170, 2008. 8.3

[Cla88] Kenneth L. Clarkson. A Randomized Algorithm for Closest-Point Queries. *SIAM J. Comput.*, 7(4):830–847, 1988. 4.2

[CO08] Frédéric Chazal and Steve Y. Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proceedings of the 24th ACM Symposium on Computational Geometry*, pages 232–241, 2008. 6.3, 6.4, 6.4.3, 7.2, 7.8

[CP03] Siu-Wing Cheng and Sheung-Hung Poon. Graded Conforming Delaunay Tetrahedralization with Bounded Radius-Edge Ratio. In *Proceedings of the Fourteenth Annual Symposium on Discrete Algorithms*, pages 295–304, Baltimore, Maryland, January 2003. Society for Industrial and Applied Mathematics. 3.5

[CSdVY04] David Cohen-Steiner, Éric Colin de Verdière, and Mariette Yvinec. Conforming Delaunay triangulations in 3D. *Comput. Geom*, 28(2-3):217–233, 2004. 3.5

[CSEH05] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Proceedings of the 21st ACM Symposium on Computational Geometry*, pages 263–271, 2005. 6.3, 6.5, 6.5.1

[CSEM06] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In *Proceedings of the 22nd ACM Symposium on Computational Geometry*, pages 119–126, 2006. 6.3

[CSZ10] Gunnar Carlsson, Gurjeet Singh, and Afra Zomorodian. Computing multidimensional persistence. *Journal of Computational Geometry*, 1(1):72–100, 2010. 6.3

[CZCG04] Gunnar Carlsson, Afra Zomorodian, Anne Colling, and Leonidas J. Guibas. Persistence barcodes for shapes. In *Eurographics Symposium on Geometry Processing*, 2004. 7.7

[dB95] Mark de Berg. Linear size binary space partitions for fat objects. In *ESA: European Symposium on Algorithms*, pages 252–263, 1995. 5.2

[dB05] Mark de Berg. Vertical ray shooting for fat objects. In *Proceedings of the 21st ACM Symposium on Computational Geometry*, pages 288–295, 2005. 5.2

[Dey07] T. K. Dey. *Curve and Surface Reconstruction : Algorithms with Mathematical Analysis*. Cambridge University Press, 2007. 4.15

[DGK99] Christian A. Duncan, Michael T. Goodrich, and Stephen Kobourov. Balanced aspect ratio trees: combining the advantages of k-d trees and octrees. In *SODA: ACM-SIAM Symposium on Discrete Algorithms*, pages 300–309, 1999. 5.2

[DI04] Mirella Damian-Iordache. Exact and approximation algorithms for computing optimal fat decompositions. *Comput. Geom. Theory Appl.*, 2004. 5.2

[dS08] Vin de Silva. A weak characterisation of the Delaunay triangulation. *Geometriae Dedicata*, 135(1):39–64, 2008. 7.2

[dSC04] V. de Silva and G. Carlsson. Topological estimation using witness complexes. In *Proc. Sympos. Point-Based Graphics*, pages 157–166, 2004. 7.2

[dSG07a] Vin de Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algorithmic & Geometric Topology*, 7:339–358, 2007. 1.2, 8.5

[dSG07b] Vin de Silva and Robert Ghrist. Homological sensor networks. *Notices Amer. Math. Soc.*, 54(1):10–17, 2007. 1.2

[dST02] Vin de Silva and Joshua B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *NIPS*, 2002. 8.3

[Dwy91] Rex A. Dwyer. Higher-Dimensional Voronoi Diagrams in Linear Expected Time. *Discrete & Computational Geometry*, 6(4):343–367, 1991. 5.2

[Ede95] Herbert Edelsbrunner. The union of balls and its dual shape. *Discrete & Computational Geometry*, 13:415–440, 1995. 6.3, 7.2

[Ede01] Herbert Edelsbrunner. *Geometry and Topology for Mesh Generation*. Cambridge University Press, 2001. 2.3

[EG02] Herbert Edelsbrunner and Damrong Guoy. An experimental study of sliver exudation. *Eng. Comput. (Lond.)*, 18(3):229–240, 2002. 1.1

[EH09] Herbert Edelsbrunner and John L. Harer. *Computational Topology: An Introduction*. Amer. Math. Soc., 2009. 6.2.2

[ELM+00] Herbert Edelsbrunner, Xiang-Yang Li, Gary L. Miller, Andreas Stathopoulos, Dafna Talmor, Shang-Hua Teng, Alper Üngör, and Noel Walkington. Smoothing and cleaning up slivers. In *STOC*, pages 273–277, 2000. 1.1

[ELZ02] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 4(28):511–533, 2002. 1.2, 6.3, 6.3, 6.3

[Eri01] Jeff Erickson. Nice point sets can have nasty Delaunay triangulations. In *Proceedings of the 17th ACM Symposium on Computational Geometry*, pages 96–105, 2001. 5.2, 7.2

[ES52] S. Eilenberg and N. Steenrod. *Foundations of Algebraic Topology*. Princeton University Press, 1952. 7.2

[GMM11] Leonidas J. Guibas, Quentin Mérigot, and Dmitriy Morozov. Witnessed k-distance. In *Proceedings of the 27th ACM Symposium on Computational Geometry*, pages 57–64, 2011. 8.4

[GO07] Leonidas J. Guibas and Steve Y. Oudot. Reconstruction using witness complexes. In *Proceedings 18th ACM-SIAM Symposium: Discrete Algorithms*, pages 1076–1085, 2007. 7.2

[Gon85] Teofilo F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theor. Comput. Sci.*, 38:293–306, 1985. 4.5

[Gra10] Chris Gray. Nearest-neighbor queries with well-spaced points. *International Symposium on Voronoi Diagrams in Science and Engineering*, 0:42–49, 2010. 5.2, 5.3, 1

[Hat01] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2001. 6.2.2, 6.2.4, 1, 2, 6.4.3

[Hen79] Michael Henle. *A Combinatorial Introduction to Topology*. Dover, 1979. 6.2.2

[HMOS10] Benoît Hudson, Gary L. Miller, Steve Y. Oudot, and Donald R. Sheehy. Topological inference via meshing. In *Proceedings of the 26th ACM Symposium on Computational Geometry*, pages 277–286, 2010. 4.15, 8.1

[HMP06] Benoît Hudson, Gary Miller, and Todd Phillips. Sparse Voronoi Refinement. In *Proceedings of the 15th International Meshing Roundtable*, pages 339–356, Birmingham, Alabama, 2006. Long version available as Carnegie Mellon University Technical Report CMU-CS-06-132. 1.1, 3.3, 3.7, 4.2, 4.4, 4.4, 4.8, 4.9, 4.11, 5.1

[HMP07] Benoît Hudson, Gary L. Miller, and Todd Phillips. Sparse Parallel Delaunay Refinement. In *19th Annual ACM Symposium on Parallelism in Algorithms and Architectures*, pages 339–347, San Diego, June 2007. 1.1, 4.2, 4.15

[HMPS09] Benoît Hudson, Gary L. Miller, Todd Phillips, and Donald R. Sheehy. Size complexity of volume meshes vs. surface meshes. In *SODA: ACM-SIAM Symposium on Discrete Algorithms*, 2009. 2.4

[HPM06] Sariel Har-Peled and Manor Mendel. Fast construction of nets in low dimensional metrics, and their applications. *SIAM Journal on Computing*, 35(5):1148–1184, 2006. 4.3

[HPÜ05] Sariel Har-Peled and Alper Üngör. A Time-Optimal Delaunay Refinement Algorithm in Two Dimensions. In *Proceedings of the 21st ACM Symposium on Computational Geometry*, pages 228–236, 2005. 4.2

[Hud07] Benoît Hudson. *Dynamic Mesh Refinement*. PhD thesis, Carnegie Mellon University, 2007. 3.5

[KC10] Jennifer Kloke and Gunnar Carlsson. Topological de-noising: Strengthening the topological signal. 2010. 8.2, 8.3, 8.4

[Kle80] Victor Klee. On the complexity of d-dimensional Voronoi diagrams. *Archiv der Mathematik*, 34:75–80, 1980. 5.2

[Lab06] François Labelle. Sliver Removal by Lattice Refinement. In *Proceedings of the 22nd ACM Symposium on Computational Geometry*. Association for Computing Machinery, June 2006. 1.1

[Li00] Xiang-Yang Li. Spacing Control and Sliver-Free Delaunay Mesh. In *Ninth International Meshing Roundtable*, pages 295–306, New Orleans, Louisiana, October 2000. 1.1

[Li03] Xiang-Yang Li. Generating well-shaped d-dimensional Delaunay meshes. *Theor. Comput. Sci.*, 296(1):145–165, 2003. 1.1

[LT01] Xiang-Yang Li and Shang-Hua Teng. Generating well-shaped Delaunay meshed in 3D. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, pages 28–37. ACM Press, 2001. 1.1

[Mar04] David J. Marchette. *Random Graphs for Statistical Pattern Recognition*. Wiley Interscience Publ., 2004. 8.3

[Mat02] Jiří Matoušek. *Lectures on Discrete Geometry*. Springer-Verlag, 2002. 3.5, 4.5

[Mér10] Quentin Mérigot. Détection de structure géométrique dans les nuages de points. Thesis tel-00443038, Université de Nice Sophia Antiopolis, 2010. 8.4

[MM99] Songrit Maneewongvatana and David M. Mount. It's okay to be skinny, if your friends are fat. In *Center for Geometric Computing 4th Annual Workshop on Computational Geometry*, 1999. 5.2, 5.4

[MMS11] Nikola Milosavljevic, Dmitriy Morozov, and Primoz Skraba. Zigzag persistent homology in matrix multiplication time. In *Proceedings of the 27th ACM Symposium on Computational Geometry*, 2011. 6.3

[MPS08] Gary L. Miller, Todd Phillips, and Donald R. Sheehy. Linear-size meshes. In *CCCG: Canadian Conference in Computational Geometry*, pages 175–178, 2008. 4.2, 4.8, 4.13, 4.15

[MTTW95] Gary L. Miller, Dafna Talmor, Shang-Hua Teng, and Noel Walkington. A Delaunay based numerical method for three dimensions: generation, formulation, and partition. In *Proceedings of the 27th Annual ACM Symposium on Theory of Computing*, pages 683–692, 1995. 1.1

[MTTW99] Gary L. Miller, Dafna Talmor, Shang-Hua Teng, and Noel Walkington. On the radius-edge condition in the control volume method. *SIAM J. on Numerical Analysis*, 36(6):1690–1708, 1999. 1.1, 3.2, 3.7, 4.8, 4.11, 5.1, 5.2, 5.3.1

[MV00] Scott A. Mitchell and Stephen A. Vavasis. Quality mesh generation in higher dimensions. *SIAM J. Comput.*, 29(4):1334–1370 (electronic), 2000. 4.2

[NSW08a] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1-3):419–441, 2008. 6.3, 8.3

[NSW08b] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. A topological view of unsupervised learning from noisy data. 2008. 8.2

[OvdS94] Mark H. Overmars and A. Frank van der Stappen. Range searching and point location among fat objects. *J. Algorithms*, 21:629–656, 1994. 5.2

[Phi09] Todd Phillips. *Efficient Mesh Generation for Piecewise Linear Complexes*. PhD thesis, Carnegie Mellon University, 2009. 3.5

[PW04] Steven E. Pav and Noel J. Walkington. Robust Three Dimensional Delaunay Refinement. In *Thirteenth International Meshing Roundtable*, pages 145–156, Williamsburg, Virginia, September 2004. Sandia National Laboratories. 3.5

[RS00] Sam T. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000. 8.3

[Rup95] Jim Ruppert. A Delaunay refinement algorithm for quality 2-dimensional mesh generation. *J. Algorithms*, 18(3):548–585, 1995. 1.1, 3.1, 3.3, 4.9

[RW08] Alexander Rand and Noel Walkington. 3d Delaunay refinement of sharp domains without a local feature size oracle. In *17th Internation Meshing Roundtable.*, 2008. 3.5

[Sei87] Raimund Seidel. On the number of faces in higher-dimensional Voronoi diagrams. In *Proceedings of the 3rd Annual Symposium on Computational Geometry*, pages 181–185, 1987. 5.2

[SF73] Gilbert Strang and George J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall, Englewood Cliffs, New Jersey, 1973. 1.1

[She97] Jonathan Richard Shewchuk. *Delaunay Refinement Mesh Generation*. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, May 1997. Available as Technical Report CMU-CS-97-137. 3.3

[She02a] Jonathan Richard Shewchuk. Delaunay Refinement Algorithms for Triangular Mesh Generation. *Computational Geometry: Theory and Applications*, 22(1–3):21–74, May 2002. 3.3

[She02b] Jonathan Richard Shewchuk. What Is a Good Linear Element? Interpolation, Conditioning, and Quality Measures. In *Eleventh International Meshing Roundtable*, pages 115–126, Ithaca, New York, September 2002. Sandia National Laboratories. 1.1

[SMI+08] Gurjeet Singh, Facundo Mémoli, Tigran Ishkhanov, Guillermo Sapiro, Gunnar Carlsson, and Dario L. Ringach. Topological analysis of population activity in visual cortex. *Journal of Vision*, 8(8):1–18, 2008. 1.2

[TdS00] Joshua B. Tenenbaum and Vin de Silva. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000. 8.3

[Üng09] Alper Üngör. Off-centers: A new type of Steiner points for computing size-optimal quality-guaranteed Delaunay triangulations. *Comput. Geom.*, 42(2):109–118, 2009. 3.3, 7.8

[vdSHO93] A. Frank van der Stappen, Dan Halperin, and Mark Overmars. The complexity of the free space for a robot moving amidst fat obstacles. *Comput. Geom. Theory Appl.*, 3:353–373, 1993. 5.2, 5.3

[vdSO94] A. Frank van der Stappen and Mark H. Overmars. Motion planning amidst fat objects. In *Proceedings of the 10th ACM Symposium on Computational Geometry*, pages 31–40, 1994. 5.2

[Vie27] L. Vietoris. Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen. *Mathematische Annalen*, 97(1):454–472, 1927. 7.2

[Yao81] Andrew Chi-Chih Yao. A lower bound to finding convex hulls. *J. ACM*, 28:780–787, October 1981. 4.2

[ZC05] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005. 1.2, 6.3, 6.3, 6.4, 6.4.2

[Zie95] Günter M. Ziegler. *Lectures on Polytopes*. Springer-Verlag, 1995. 2.1

[Zom09] Afra Zomorodian. *Topology for Computing*. Cambridge Univ. Press, 2009. 6.2.2

[ZZZ] Plex. PLEX 2.5. See http://comptop.stanford.edu/programs/plex.html. 7.7