CURVE AND SURFACE RECONSTRUCTION FROM NOISY SAMPLES

by

SHEUNG-HUNG POON

A Thesis Submitted to
The Hong Kong University of Science and Technology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
in Computer Science

May 2004, Hong Kong

Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the Hong Kong University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

SHEUNG-HUNG POON

CURVE AND SURFACE RECONSTRUCTION FROM NOISY SAMPLES

by

SHEUNG-HUNG POON

This is to certify that I have examined the above Ph.D. thesis
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by
the thesis examination committee have been made.

DR. SIU-WING CHENG, THESIS SUPERVISOR

PROF. LIONEL M. NI, HEAD OF DEPARTMENT

Department of Computer Science

22 May 2004

To My Wife

Yuki

and Family

for their Patience and Understanding

ACKNOWLEDGMENTS

Special thanks go to my thesis supervisor Dr. Siu-Wing Cheng for his brilliant guidance of my work and great patience to lead me to be a researcher. Collaboration with him results in widening and enriching my styles of thinking, correcting my directions to solve a problem, and improving my communication and writing skills. Thanks also go to the members of my thesis examination committee for taking the time to read my thesis and offer important comments. Moreover, thanks also go to my colleagues and my friends in the Theoretical Computer Science Group in our department.

Thanks to the system staff in our major department. They are very helpful and their technical support are excellent.

Last but not least, I sincerely thank my family for their support, and great encouragement for my studies of the doctoral degree. Without them, this thesis would never come to its existence.

TABLE OF CONTENTS

LIST OF FIGURES

CURVE AND SURFACE RECONSTRUCTION FROM NOISY SAMPLES

by

SHEUNG-HUNG POON

Department of Computer Science

The Hong Kong University of Science and Technology

ABSTRACT

Reconstructing an unknown curve or surface from sample points is an important task in geometric modeling applications. Sample points obtained from real applications are usually noisy. For example, when data sets are obtained by scanning images in the plane or images in three dimensions. In computer graphics, many curve and surface reconstruction algorithms have been developed. However, their common drawback is the lack of theoretical guarantees on the quality of the reconstruction. This motivates computational geometers to propose algorithms that return provably faithful reconstructions. Algorithms of this type are known when there is no noise in the input. This leaves the problem of noise handling open. We propose a probabilistic noise model for the curve reconstruction problem. Based on this model, we design a curve reconstruction algorithm for noisy input points. The reconstruction is faithful with probability approaching 1 as the sampling density increases. Then we extend our approach to surface reconstruction from noisy input points. Not

only do we improve the algorithm to make it run faster, we also make the noise model deterministic which extends its applicability and simplifies the analysis of the algorithm. We show that the surface reconstructed is faithful if the input points satisfy the deterministic noise model.

CHAPTER 1

INTRODUCTION

The problems of reconstructing curves and surfaces from sample points has been studied extensively in computer graphics, geometric modeling, image processing, computer vision, and computational geometry. The input consists of sample points from the unknown smooth closed curve or surface S. The problem calls for computing a polygonal curve / surface that is provably *faithful*. That means, when the sampling density is sufficiently large, the reconstructed polygonal curve / surface should be homeomorphic to S and the approximation error should approach zero. There are two types of approximation error. First, the Hausdorff distance between the reconstruction and S. Second, the differences between the normals of the reconstruction and that of S. Both types of error should approach zero as the sampling density increases.

Practical sample point sets may be noisy. For example, two-dimensional images obtained by scanning, three-dimensional medical data, and three-dimensional scanned data points obtained by 3D scanners. In computer graphics and computer vision, a lot of reconstruction algorithms have been proposed to handle noise. However, their common drawback is the lack of theoretical guarantees even when the input samples are very dense. This motivates computational geometers to design reconstruction algorithms that guarantee faithfulness. Several algorithms have been proposed for noiseless sample points. Less is known about the faithful surface reconstruction problem from noisy samples.

The noisy samples are typically classified into two types. The first type are samples that cluster around the curve / surface S but they generally do not lie on it. The second type are outliers that lie relatively far from the curve / surface S. In this thesis, we only consider the handling of noise of the first type with theoretical guarantees. It is more difficult to handle the noise of second type with theoretical

guarantees, which is out of scope of this thesis. There are known filtering methods for removing outliers [9, 27, 47, 48, 51, 52].

We summarize our main contributions in this thesis here. We first propose a probabilistic noise model for the curve reconstruction problem. Based on this model, we design a curve reconstruction algorithm for the noisy input. The reconstruction is faithful with probability approaching 1 as the sampling density increases. Then we extend our approach to perform surface reconstruction from the noisy input points. We are able to make the noise model deterministic so that the analysis of the faithfulness proof is simplified. We prove that the surface reconstructed is faithful to the original surface if the input points satisfy the deterministic noise model. Both of our curve and surface reconstruction algorithms follow the same framework. Our algorithms first construct a set of center points from the given noisy sample points. The center points are provably much less noisy than the input points. A subset of these center points is then selected for the final surface reconstruction. This step can be done using any existing faithful reconstruction algorithm for noiseless samples. We choose the NN-crust algorithm [16] for reconstructing curves and the cocone algorithm [5] for reconstructing surfaces, respectively.

We give a brief survey on the related work for curve and surface reconstruction in Sections 1.1 and 1.2, respectively. Then we give an outline of this thesis in Section 1.3.

1.1 Related work for curve reconstruction

Several curve reconstruction algorithms have been proposed in the geometric modeling and image processing literature that achieve good experimental results. Fang and Gossard [24] proposed to fit a deformable curve by minimizing some spring energy function. Dedieu and Favardin [13] described a method to order and connect sample points on an unknown 2D curve. Taubin and Ronfard [50] proposed to construct a mesh covering the sample points and then extract a polygonal curve that fits the sample points. Pottmann and Randrup [43] used a pixel-based technique to thin an input point cloud to a curve. This image thinning technique can handle

noise, but it is difficult to come up with an appropriate pixel size. Goshtasby [31] obtained a reconstruction by tracing points that locally maximize a certain inverse distance function involving the noisy sample points. The traced points form the reconstruction. Lee [37] proposed a variant of the moving least-squares method by Levin [38, 39]. Using a weighted regression, a new point is computed for each noisy sample point such that the new points cluster around some curve. Then the new points are decimated to produce a reconstruction. Although good experimental results are obtained with the above methods, there was no guarantee on the faithfulness of the reconstruction.

Algorithms in [10, 44, 45, 46] tried to select a subset of Delaunay edges to approximate the original curve. Edelsbrunner, Kirkpatrick, and Seidel [21] proposed the *α-shape* for 2D shape reconstruction. Their algorithm tends to work well for sample points with uniform distribution on the domain. Amenta, Bern, and Eppstein [3] obtained the first result for curve reconstruction problem from noiseless samples with theoretical guarantees. They proposed a *2D crust* algorithm whose output is provably faithful if the input satisfies the *ϵ-sampling* condition for any $\epsilon < 0.252$. For each point x on S, the *local feature size* $f(x)$ at x is defined as the distance from x to the medial axis of S. For $0 < \epsilon < 1$, a set S of samples is an *ϵ-sampling* of S if for any point $x \in S$, there exists a sample $s \in S$ such that $d(s, x) \leq \epsilon \cdot f(x)$ [3]. Note that the definition for local feature size applies also to higher dimensions. The 2D crust algorithm invokes the computation of a Voronoi diagram or Delaunay triangulation twice. Gold and Snoeyink [30] presented a simpler algorithm that invokes the computation of Delaunay triangulation only once. Later, Dey and Kumar [16] proposed an even simpler algorithm called *NN-crust*. They showed that the output of their algorithm is faithful to the original curve if the input sample points satisfy the ϵ-sampling condition for any $\epsilon \leq 1/3$ in [16].

Dey, Mehlhorn, and Ramos [17] proposed an algorithm to handle curves with endpoints with guarantees. Dey and Wenger [18, 19] presented an algorithm to handle sharp corners without guarantees; whereas Funke and Ramos [25] can do that with guarantees. Giesen [28] discovered that the traveling salesperson tour through the samples is a faithful reconstruction. Althaus and Mehlhorn [1] showed

that such a traveling salesperson tour can be constructed in polynomial time.

1.2 Related work for surface reconstruction

The surface reconstruction problem has been studied extensively in computer graphics and computer vision. The first and widely known reconstruction algorithm in computer graphics is the work of Hoppe et al. [34, 36, 35]. Their work performs quite well in practice, and allows the presence of noise. Later, Curless and Levoy [12] presented a volumetric method to reconstruct surface from range images. Medioni and Tang [40, 49] applied a technique called tensor voting to infer the surface and features on it. All these algorithms can also handle noise in the input, and appear to be quite successful in practice. However, they do not have theoretical guarantees.

Some faithful reconstruction algorithms were proposed in computational geometry for the simple case that sample points are noiseless. Edelsbrunner and Mucke [22] used the α-shape to construct the surface, and Edelsbrunner [20] proposed another surface reconstruction algorithm by wrapping the sample points. The main drawback is that they perform well only for uniform sample set. Amenta and Bern [2, 4] extended the approach of Voronoi filtering to do surface reconstruction. Later, Amenta, Choi, Dey, and Leekha [5] presented a simpler algorithm with a proof that the reconstruction is homeomorphic to the original surface. Amenta et al. [6, 7] gave another faithful reconstruction algorithm, which is called power crust. One merit of the power crust algorithm is that it always output a water-tight surface regardless of the sampling density of the input data set. Another algorithm given by Boissonnat and Cazals [8] reconstructs a smooth implicit surface interpolating the sample points. One can then mesh the implicit surface to obtain a mesh of the desired resolution.

As we mentioned before, noise often arises in practical sample point set. Recently, Mitra and Nguyen [42] presented methods to estimate surface normals from point clouds, but did not proceed to finally reconstruct the surface for the sample points. Very recently and independently, Dey and Goswami [15] proposed a surface reconstruction algorithm from noisy samples. There are several major differences

between their work and ours in this thesis. In their noise model, the noise amplitude is proportional to the local feature size and inversely proportional to the sampling density; whereas in our model, it is an independent constant. Their algorithm selects a set of "boundary" noisy sample points for reconstruction; whereas our algorithm tries to estimate some center points close to the original surface for reconstruction. Finally, they gave an implementation for their algorithm and had some promising results, but we havn't implemented our algorithm.

1.3 Thesis outline

We gave an introduction for the curve and surface reconstruction problem in the current chapter (Chapter 1).

In Chapter 2, we propose a particular probabilistic model for noisy samples for smooth curves on the plane. And in Chapter 3, provided that the input sample points follows the probabilistic noise model, we prove that our reconstructed curve is faithful to the original curve with probability approaching 1 as the number of samples increases. The novelty of our algorithm is a method to cluster samples so that each cluster comes from a relatively flat portion of S. This allows us to further estimate points that lie close to S.

In Chapter 4, we propose a deterministic noise model for noisy samples for smooth surfaces in three dimensions, and justify it. Making use of this deterministic model, we can have a simpler and cleaner analysis for the faithfulness proof of our surface reconstruction algorithm in Chapter 5. Our surface reconstruction algorithm is a generalized version of the algorithm in Chapter 3.

And finally in Chapter 6, we conclude this thesis and propose some future research problems.

CHAPTER 2

PROBABILISTIC NOISE MODEL FOR CURVES

In this and next chapters, we only consider the given sample points are drawn from an unknown closed smooth curve S on the plane. Suppose P is the given set of n noisy sample points from S. As we mentioned in the introduction chapter, we only consider noise of the first type, which are sample points cluster around and close to S. We denote the maximum noise amplitude away from S by δ. We assume that $\min_{x \in S} f(x) = 1$ for convenience. So for any $x \in S$, $f(x) \geq 1$.

In this chapter, we present a probabilistic noise model for a curve (Section 2.1). We then introduce the basic notations and some basic geometric lemmas in Section 2.2. Finally in Section 2.3, we will look at some consequent technical lemmas due to this model, which are needed later in the analysis of our curve reconstruction algorithm in Chapter 3. Note that the proofs of some technial lemmas are given in the appendix.

2.1 Sampling and noise model

We use probabilistic sampling to model the noise. A sample is generated by drawing a point from S followed by randomly perturbing the point in the normal direction. In a sense, it models the location of points on the curve by an input device, followed by perturbation due to noise. Let $L = \int_S \frac{1}{f(x)} dx$. The drawing of points from S follows the probability density function $\frac{1}{L \cdot f(x)}$. That is, the probability of drawing a point from a curve segment η is equal to $\int_\eta \frac{1}{f(x)} dx$ divided by L. This is known as the *locally uniform distribution*. The distribution of each sample is independently identical.

A point p drawn from S is perturbed in the normal direction. The perturbation is uniformly distributed within an interval that has p as the midpoint, width 2δ, and

aligns with the normal direction at p. Note that the noise amplitude δ remains fixed regardless of the number of points drawn from S. Although the noise perturbation is restrictive, it isolates the effect of noise from the sampling distribution which allows an initial study of noise handling. It seems necessary that δ is less than 1. Otherwise, as the minimum local feature size is 1, the perturbed points from different parts of S will mix up at some place and it seems very difficult to estimate the unknown curve S around that neighborhood. For our analysis to work, we assume that $\delta \leq 1/(25\rho^2)$ where $\rho \geq 5$ is a constant chosen a priori by our algorithm. We emphasize that the value of δ is unknown to our algorithm.

One may consider other sampling distributions. A more restrictive model is the *uniform distribution*, in which the probability of drawing a point from a curve segment η is equal to $\frac{\text{length}(\eta)}{\text{length}(S)}$. This model is attractive because it is natural to sample in a uniform fashion in the absence of any information about the local feature sizes. Despite the apparent difference, the locally uniform distribution is strongly related to the uniform distribution which can be seen as follows. When η is short, the Lipschitz property of the local feature sizes implies that the probability of drawing a point from η in the locally uniform model is $\Theta(\frac{\int_\eta dx}{L \cdot f(c)})$ for any point $c \in \eta$. This is equivalent to $\Theta(\frac{\text{length}(\eta)}{L \cdot f(c)})$. If we treat L and length(S) as intrinsic constants for S, the probabilities of sampling in the locally uniform distribution and the uniform distribution differ only by a factor of local feature size. Thus our analysis for the locally uniform distribution can be adapted easily for the uniform distribution case, basically by slashing off a factor of local feature size. In particular, the reconstruction is faithful with probability at least $1 - O(n^{-\Omega(\ln^\omega n - 1)})$ instead of $1 - O(n^{-\Omega(\frac{\ln^\omega n}{f_{\max}} - 1)})$.

Our algorithm and analysis do not make use of any estimation of local feature sizes. This is demonstrated by the fact that our analysis can be adapted to the uniform distribution case as briefly explained above. Our algorithm constructs a small neighborhood around each noisy sample, and from this small neighborhood, one can extract upper and lower bounds on the local feature size. However, the two bounds differ by a factor that tends to infinity as the sampling density increases. So the small neighborhood does not offer any reliable estimation of the local feature size. (We will elaborate on this point when we describe our algorithm.) In fact,

we do not know how to obtain such estimation in the presence of noise, without effectively solving the reconstruction problem first. After solving the reconstruction problem, one may possibly estimate the local feature sizes using the Voronoi diagram of the reconstruction as an approximation of the medial axis. This is beyond the scope of this paper though.

2.2 Preliminaries

We call the bounded region enclosed by S the *inside* of S and the unbounded region the *outside* of S. For $0 < \alpha \leq \delta$, S_α^+ (resp., S_α^-) is the curve that passes through the points q outside (resp., inside) S such that $d(q, \tilde{q}) = \alpha$. We use S_α to mean S_α^+ or S_α^- when it is unimportant to distinguish between inside and outside. S can be visualized as the boundary of the union of the medial disks enclosed by S. If we increase the radii of all such medial disks by α, S_α^+ is the boundary of the union of the expanded disks. S_α^- has a similar interpretation after decreasing the radii of all such medial disks by α. It follows that S and S_α have the same medial axis.

The *normal segment* at a point $p \in S$ is the line segment consisting of the points q on the normal of S at p such that $d(p, q) \leq \delta$. Given two points x and y on S, we use $S(x, y)$ to denote the curved segment traversed from x to y in clockwise direction. We use $|S(x, y)|$ to denote the length of $S(x, y)$.

The following are some technical lemmas on some geometric properties of S_α. Their proofs can be found in the appendix. Lemma 2.2.1 lower bounds the radius of the tangent disk at any point on S_α. Lemma 2.2.2 shows that a small neighborhood of a point p on S_α is flat enough to fit inside a double cone at p with small aperture. Lemma 2.2.3 proves the small normal variation between two nearby points on S_α.

Lemma 2.2.1 *Any point p on S_α has two tangent disks with radii $f(\tilde{p}) - \alpha$ whose interior do not intersect S_α.*

For each point p on S_α, take the double cone of points q such that pq makes an angle $(\pi - \theta)/2$ or less with the support line of the normal at p. We denote the

complement of this double cone by $cocone(p, \theta)$. Note that $cocone(p, \theta)$ is a double cone with apex p and angle θ.

Lemma 2.2.2 *Let p be a point on S_α. Let D be a disk centered at p with radius less than $2(1-\alpha)f(\tilde{p})$.*

(i) For any point $q \in S_\alpha \cap D$, the distance of q from the tangent at p is at most $\frac{d(p,q)^2}{2(1-\alpha)f(\tilde{p})}$.

(ii) $S_\alpha \cap D \subseteq cocone(p, 2\sin^{-1}\frac{\text{radius}(D)}{2(1-\alpha)f(\tilde{p})})$.

Lemma 2.2.3 *Let p be a point on S_α. Let D be a disk centered at p with radius at most $\frac{(1-\alpha)f(\tilde{p})}{4}$. For any point $u \in S_\alpha \cap D$, the acute angle between the normals at p and u is at most $2\sin^{-1}\frac{d(p,u)}{(1-\alpha)f(\tilde{p})} \leq 2\sin^{-1}\frac{\text{radius}(D)}{(1-\alpha)f(\tilde{p})}$.*

2.3 Decompositions

We will use two types of decompositions, *β-partition* and *β-grid*. Let $0 < \beta < 1$ be a parameter. We identify a set of *cut-points* on S as follows. We pick an arbitrary point c_0 on S as the first cut-point. Then for $i \geq 1$, we find the point c_i such that c_i lies in the interior of $S(c_{i-1}, c_0)$, $|S(c_{i-1}, c_i)| = \beta^2 f(c_{i-1})$, and $|S(c_i, c_0)| \geq \beta^2 f(c_i)$. If c_i exists, it is the next cut-point and we continue. Otherwise, we have computed all the cut-points and we stop. The β-partition is the arrangement of S_δ^+, S_δ^-, and the normal segments at the cut-points. Figure 2.1 shows an example. We call each face of the β-partition a *β-slab*. The β-partition consists of a row of slabs stabbed by S.

The cut-points for a β-grid are picked differently. We pick an arbitrary point c_0 on S as the first cut-point. Then for $i \geq 1$, we find the point c_i such that c_i lies in the interior of $S(c_{i-1}, c_0)$, $|S(c_{i-1}, c_i)| = \beta f(c_{i-1})$, and $|S(c_i, c_0)| \geq \beta f(c_i)$. If c_i exists, it is the next cut-point and we continue. Otherwise, we have computed all the cut-points and we stop. The β-grid is the arrangement of the following:

- The normal segments at the cut-points.

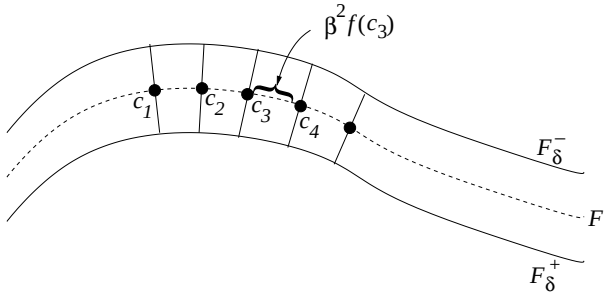

Figure 2.1: β-partition.

- S, S_δ^+, and S_δ^-.
- S_α^+ and S_α^- where $\alpha = i\beta\delta$ and i is an integer between 1 and $\lfloor 1/\beta \rfloor - 1$.

The β-grid has a grid structure. Figure 2.2 shows an example. We call each face of the β-grid a *β-cell*. There are $O(1/\beta)$ rows of cells "parallel to" S.

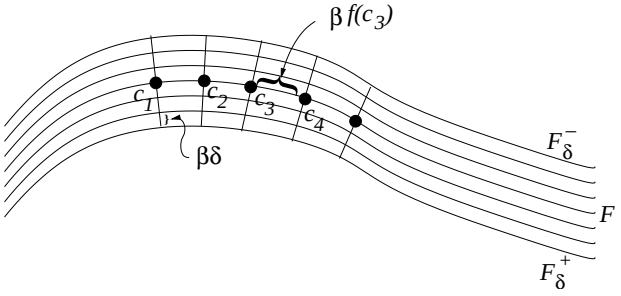

Figure 2.2: β-grid.

Given a β-partition, we claim that for every consecutive pairs of cut-points c_{i-1} and c_i, $\beta^2 f(c_{i-1}) \leq |S(c_{i-1}, c_i)| \leq 3\beta^2 f(c_{i-1})$. For almost all consecutive pairs of cut-points c_{i-1} and c_i, $|S(c_{i-1}, c_i)| = \beta^2 f(c_{i-1})$ by construction. The last pair c_k and c_0 constructed may be an exception. We know that $|S(c_k, c_0)| \geq \beta^2 f(c_k)$. When we try to place c_{k+1}, we find that $|S(c_{k+1}, c_0)| < \beta^2 f(c_{k+1})$. So $|S(c_k, c_0)| \leq \beta^2 f(c_k) + \beta^2 f(c_{k+1})$. By the Lipschitz condition, $f(c_{k+1}) \leq f(c_k) + d(c_k, c_{k+1}) \leq f(c_k) + \beta^2 f(c_k)$. Thus $|S(c_k, c_0)| \leq (2\beta^2 + \beta^4) f(c_k) \leq 3\beta^2 f(c_k)$.

Similarly, given a β-grid, we can show that for every consecutive pairs of cut-points c_{i-1} and c_i, $\beta f(c_{i-1}) \leq |S(c_{i-1}, c_i)| \leq 3\beta f(c_{i-1})$.

In Section 2.3.1, we bound the diameter of a β-cell. In Section 2.3.2, we lower

bound the width of a β-slab. In Section 2.3.3, we analyze the probabilities of some β-slabs and β-cells containing certain numbers of samples.

2.3.1 Diameter of a β-cell

We need a technical lemma before proving an upper bound on the diameter of a β-cell.

Lemma 2.3.1 *Assume that $\beta \leq 1/12$. Let p and q be two points on S_α such that $|S(\tilde{p}, \tilde{q})| \leq 3\beta f(\tilde{p})$. Then $d(p,q) \leq d(\tilde{p}, \tilde{q}) + 7\beta\delta$.*

Proof. Refer to Figure 2.3. Let r be the point $q - \tilde{q} + \tilde{p}$. Without loss of generality, assume that $\angle\tilde{p}pr \leq \angle\tilde{p}rp$. Lemma 2.2.3 implies that $\angle p\tilde{p}r \leq 2\sin^{-1} 3\beta$. Therefore,

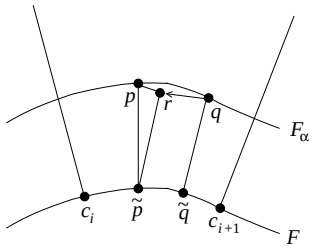

Figure 2.3: Illustration for Lemma 2.3.1.

$\angle\tilde{p}rp \geq \pi/2 - \sin^{-1} 3\beta$. By sine law, $d(p,r) = \frac{d(p,\tilde{p})\cdot\sin\angle p\tilde{p}r}{\sin\angle\tilde{p}rp} \leq \frac{\delta\sin(2\sin^{-1}3\beta)}{\cos(\sin^{-1}3\beta)}$. Note that $\sin(2\sin^{-1}3\beta) \leq 2\sin(\sin^{-1}3\beta) = 6\beta$ and since $\beta \leq 1/12$, $\cos(\sin^{-1}3\beta) \geq \cos(\sin^{-1}(1/4)) > 0.9$. So $d(p,r) \leq 6\beta\delta/(0.9) < 7\beta\delta$. By triangle inequality, we get $d(p,q) \leq d(q,r) + d(p,r) = d(\tilde{p},\tilde{q}) + d(p,r) < d(\tilde{p},\tilde{q}) + 7\beta\delta$. □

Lemma 2.3.2 *Assume that $\beta \leq 1/12$ and $\delta < 1$. Let C be any β-cell that lies between the normal segments at the cut-points c_i and c_{i+1}. Then the diameter of C is at most $14\beta f(c_i)$.*

Proof. Let s and t be two points in C. Let p be the projection of s towards $\tilde{s}$ onto a side of C. Similarly, let q be the projection of t towards $\tilde{t}$ onto the same side of

C. Note that $\tilde{p} = \tilde{s}$ and $\tilde{q} = \tilde{t}$. The triangle inequality and Lemma 2.3.1 imply that

$$\begin{aligned} d(s,t) &\leq d(p,q) + d(p,s) + d(q,t) \\ &\leq d(\tilde{p},\tilde{q}) + 7\beta\delta + d(p,s) + d(q,t). \end{aligned}$$

Since $d(\tilde{p},\tilde{q}) = d(\tilde{s},\tilde{t}) \leq 3\beta f(c_i)$ and both $d(p,s)$ and $d(q,t)$ are at most $2\beta\delta$, the diameter of C is at most $3\beta f(c_i) + 11\beta\delta \leq 14\beta f(c_i)$. □

2.3.2 Slab width

The next lemma lower bounds the width of slab in a β-partition.

Lemma 2.3.3 *Assume that $\delta \leq 1/8$ and $\beta \leq 1/6$. Let c_i and c_{i+1} be two consecutive cut-points of a β-partition. For any point on the normal segment at c_{i+1} (resp., c_i), its distance from the support line of the normal segment at c_i (resp., c_{i+1}) is at least $|S(c_i, c_{i+1})|/6$.*

Proof. Assume that the normal at c_i is vertical. Take any two points $p, q \in S_\alpha$ such that $\tilde{p} = c_i$ and $\tilde{q} = c_{i+1}$. We first bound the distance from q to the support line of the normal segment at c_i. The same approach also works for the distance from p to the support line of the normal segment at c_{i+1}.

Let r be the orthogonal projection of q onto the tangent to S_α at p. Observe that the distance of q from the support line of the normal segment at c_i is $d(p,r)$. We are to prove that $d(p,r) \geq |S(c_i, c_{i+1})|/6$. For any point $x \in S_\alpha(p,q)$, we use θ_x to denote the angle between the normals at $\tilde{x}$ and c_i. By Lemma 2.2.3, we have $\theta_x \leq 2\sin^{-1}\frac{d(c_i,\tilde{x})}{f(c_i)}$. Since $\tilde{x} \in S(c_i, c_{i+1})$, we have $d(c_i,\tilde{x}) \leq |S(c_i,\tilde{x})| \leq |S(c_i, c_{i+1})|$. Thus $\theta_x \leq 2\sin^{-1}\frac{|S(c_i,c_{i+1})|}{f(c_i)}$. By our assumption on β, $\frac{|S(c_i,c_{i+1})|}{f(c_i)} \leq 3\beta^2 \leq 1/12$. It follows that $\sin^{-1}\frac{|S(c_i,c_{i+1})|}{f(c_i)} < \frac{2|S(c_i,c_{i+1})|}{f(c_i)}$. Therefore,

$$\theta_x \leq \frac{4|S(c_i,c_{i+1})|}{f(c_i)} \tag{2.1}$$

$$\leq 12\beta^2. \tag{2.2}$$

This implies that $S_\alpha(p,q)$ is monotone along the tangent to S_α at p; otherwise, there is a point $x \in S_\alpha(p,q)$ such that $\theta_x = \pi/2 > 12\beta^2$, a contradiction. It follows that $S(c_i, c_{i+1})$ is also monotone along the tangent to S at c_i. Refer to Figure 2.4. Assume

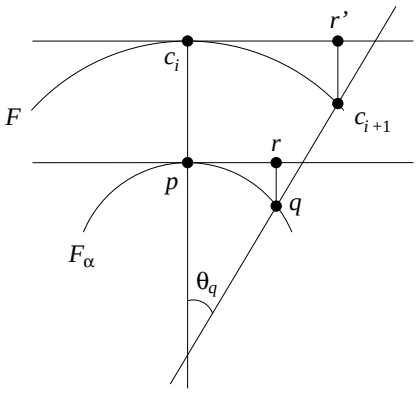

Figure 2.4: Illustration for Lemma 2.3.3.

that p lies below c_i, and q lies to the right of p. Let r' be the orthogonal projection of c_{i+1} onto the tangent to S at c_i. The monotonicity of $S(c_i, c_{i+1})$ implies that

$$d(c_i, r') = \int_{S(c_i, c_{i+1})} \cos\theta_x \, dx \overset{(2.2)}{\geq} |S(c_i, c_{i+1})| \cdot \cos(12\beta^2) > 0.8|S(c_i, c_{i+1})|,$$

as $\cos(12\beta^2) > \cos(0.5) > 0.8$. Let d be the horizontal distance between r and r'. Observe that $d = d(c_{i+1}, q) \cdot \sin\theta_q \leq \delta\theta_q$, which is at most $4\delta|S(c_i, c_{i+1})|$ by (2.1). We conclude that

$$\begin{aligned} d(p,r) &\geq d(c_i, r') - d \\ &\geq (0.8 - 4\delta)|S(c_i, c_{i+1})| \\ &\overset{\delta \leq 1/8}{>} \frac{|S(c_i, c_{i+1})|}{4}. \end{aligned}$$

This lower bounds the distance from q to the support line of the normal segment at c_i.

Let d_p be the distance from p to the support line of the normal segment at c_{i+1}. We can use the same approach to lower bound d_p. The only difference is that for any point $x \in S_\alpha(p,q)$, the angle ϕ_x between the normals at $\tilde{x}$ and c_{i+1} satisfies

$$\phi_x \leq 2 \sin^{-1} \frac{|S(c_i, c_{i+1})|}{f(c_{i+1})}.$$

Note that the denominator is $f(c_{i+1})$ instead of $f(c_i)$ in (2.1). Nevertheless, by the Lipschitz condition, $f(c_{i+1}) \geq f(c_i) - d(c_i, c_{i+1}) \geq f(c_i) - |S(c_i, c_{i+1})| \geq (1 - 3\beta^2)f(c_i)$, which is at least $11f(c_i)/12$ as $3\beta^2 \leq 1/12$. Therefore,

$$\phi_x \leq 2 \sin^{-1} \frac{12|S(c_i, c_{i+1})|}{11f(c_i)} \leq 2 \cdot \frac{24|S(c_i, c_{i+1})|}{11f(c_i)} < \frac{5|S(c_i, c_{i+1})|}{f(c_i)} \leq 15\beta^2.$$

Observe that $\phi_x \leq 15\beta^2 < \pi/2$. So $S_\alpha(p, q)$ and $S(c_i, c_{i+1})$ are monotone along the tangents to S_α at q and S at c_{i+1}, respectively. Also, $\cos\phi_x \geq \cos(15\beta^2) \geq \cos(0.5) > 0.8$. Hence, by imitating the previous derivation of the lower bound of $d(p, r)$, we obtain

$$\begin{aligned} d_p &\geq (0.8 - 5\delta)|S(c_i, c_{i+1})| \\ &\overset{\delta \leq 1/8}{>} \frac{|S(c_i, c_{i+1})|}{6}. \end{aligned}$$

□

2.3.3 Number of samples in cells and slabs

We first need a lemma that estimates the probability of a sample point lying inside certain β-cells and β-slabs.

Lemma 2.3.4 *Let $\lambda_k = \sqrt{\frac{k^2 \ln^{1+\omega} n}{n}}$ for some positive constant k. Let $r \geq 1$ be a parameter. Let C be a (λ_k/r)-slab or (λ_k/r)-cell. Let s be a sample. There exist constants κ_1 and κ_2 such that if n is so large that $\lambda_k \leq 1/6$, then $\kappa_2\lambda_k^2/r^2 \leq \Pr(s \in C) \leq \kappa_1\lambda_k^2/r^2$.*

Proof. Recall that $L = \int_S \frac{1}{f(x)}dx$. Assume that C lies between the normal segments at the cut-points c_i and c_{i+1}. We use η to denote $S(c_i, c_{i+1})$ as a short hand. By our assumption on λ_k, for any point $x \in \eta$, if C is a λ_k-cell, then $d(x, c_i) \leq 3\lambda_k f(c_i)/r \leq f(c_i)/2$; if C is a λ_k-slab, then $d(x, c_i) \leq 3\lambda_k^2 f(c_i)/r^2 \leq f(c_i)/12$. The Lipschitz condition implies that $f(c_i)/2 \leq f(x) \leq 3f(c_i)/2$. If C is a λ_k-slab, then $\Pr(s \in C) = \Pr(\tilde{s}$ lies on $\eta)$, which is $\frac{1}{L} \cdot \int_\eta \frac{1}{f(x)}dx \in [\frac{2\lambda_k^2}{3Lr^2}, \frac{6\lambda_k^2}{Lr^2}]$. If C is λ_k-cell,

then $\Pr(\tilde{s} \text{ lies on } \eta) = \frac{1}{L} \cdot \int_\eta \frac{1}{f(x)} dx \in [\frac{2\lambda_k}{3Lr}, \frac{6\lambda_k}{Lr}]$. Since $\Pr(s \in C \mid \tilde{s} \text{ lies on } \eta) \in [\frac{\lambda_k \delta}{2\delta r}, \frac{2\lambda_k \delta}{2\delta r}] = [\frac{\lambda_k}{2r}, \frac{\lambda_k}{r}]$, $\Pr(s \in C) \in [\frac{\lambda_k^2}{3Lr^2}, \frac{6\lambda_k^2}{Lr^2}]$. □

The following Chernoff bound [32] will be needed.

Lemma 2.3.5 *Let the random variables $X_1, X_2, \ldots, X_n$ be independent, with $0 \leq X_i \leq 1$ for each i. Let $S_n = \sum_{i=1}^n X_i$, and let $E(S_n)$ be the expected value of S_n. Then for any $\sigma > 0$, $\Pr(S_n \leq (1-\sigma)E(S_n)) \leq \exp(-\frac{\sigma^2 E(S_n)}{2})$, and $\Pr(S_n \geq (1+\sigma)E(S_n)) \leq \exp(-\frac{\sigma^2 E(S_n)}{2(1+\sigma/3)})$.*

We are ready to analyze the probabilities of some β-slabs and β-cells containing certain numbers of samples.

Lemma 2.3.6 *Let $\lambda_k = \sqrt{\frac{k^2 \ln^{1+\omega} n}{n}}$ for some positive constant k. Let $r \geq 1$ be a parameter. Let C be a (λ_k/r)-slab or (λ_k/r)-cell. Let κ_1 and κ_2 be the constants in Lemma 2.3.4. Whenever n is so large that $\lambda_k \leq 1/6$, the following hold.*

(i) C is non-empty with probability at least $1 - n^{-\Omega(\ln^\omega n / r^2)}$.

(ii) Assume that $r = 1$. For any constant $\kappa > \kappa_1 k^2$, the number of samples in C is at most $\kappa \ln^{1+\omega} n$ with probability at least $1 - n^{-\Omega(\ln^\omega n)}$.

(iii) Assume that $r = 1$. For any constant $\kappa < \kappa_2 k^2$, the number of samples in C is at least $\kappa \ln^{1+\omega} n$ with probability at least $1 - n^{-\Omega(\ln^\omega n)}$.

Proof. Let $X_i (i = 1, \ldots, n)$ be a random binomial variable taking value 1 if the sample point s_i is inside C, and value 0 otherwise. Let $S_n = \sum_{i=1}^n X_i$. Then $E(S_n) = \sum_{i=1}^n E(X_i) = n \cdot \Pr(s_i \in C)$. This implies that

$$E(S_n) \leq \frac{\kappa_1 n \lambda_k^2}{r^2} = \frac{\kappa_1 k^2 \ln^{1+\omega} n}{r^2}, \qquad E(S_n) \geq \frac{\kappa_2 n \lambda_k^2}{r^2} = \frac{\kappa_2 k^2 \ln^{1+\omega} n}{r^2}.$$

By Lemma 2.3.5,

$$\begin{aligned} \Pr(S_n \leq 0) &= \Pr(S_n \leq (1-1)E(S_n)) \\ &\leq \exp(-\frac{E(S_n)}{2}) \\ &\leq \exp(-\Omega(\frac{\ln^{1+\omega} n}{r^2})). \end{aligned}$$

Consider (ii). Let $\sigma = \frac{\kappa}{\kappa_1 k^2} - 1 > 0$. Since $r = 1$, we have

$$\kappa \ln^{1+\omega} n = \kappa_1 n \lambda_k^2 (1+\sigma) \geq (1+\sigma) E(S_n).$$

By Lemma 2.3.5,

$$\begin{aligned} \Pr(S_n > \kappa \ln^{1+\omega} n) &\leq \Pr(S_n > (1+\sigma)E(S_n)) \\ &\leq \exp(-\frac{\sigma^2 E(S_n)}{2+2\sigma/3}) \\ &= \exp(-\Omega(\ln^{1+\omega} n)). \end{aligned}$$

Consider (iii). Let $\sigma = 1 - \frac{\kappa}{\kappa_2 k^2} > 0$. Since $r = 1$, we have

$$\kappa \ln^{1+\omega} n = \kappa_2 n \lambda_k^2 (1-\sigma) \leq (1-\sigma) E(S_n).$$

By Lemma 2.3.5,

$$\begin{aligned} \Pr(S_n < \kappa \ln^{1+\omega} n) &\leq \Pr(S_n < (1-\sigma)E(S_n)) \\ &\leq \exp(-\frac{\sigma^2 E(S_n)}{2}) \\ &= \exp(-\Omega(\ln^{1+\omega} n)). \end{aligned}$$

□

2.4 Appendix

Proof of Lemma 2.2.1

Let M_α be the medial disk of S_α touching a point $p \in S_\alpha$. By the definition of S_α, there is a medial disk M of S touching $\tilde{p}$ such that M and M_α have the same center. Moreover, $\text{radius}(M_\alpha) = \text{radius}(M) - \alpha \geq f(\tilde{p}) - \alpha$. □

Proof of Lemma 2.2.2

Assume that the tangent at p is horizontal. Consider (i). Refer to Figure 2.5(a). Let B be the tangent disk at p that lies above p and has center x and radius $(1-\alpha)f(\tilde{p})$. Let C be the circle centered at p with radius $d(p,q)$. Since $d(p,q) < 2(1-\alpha)f(\tilde{p})$, C crosses B. Let r be a point in $C \cap \partial B$. Let d be the distance of r from the tangent at p. By Lemma 2.2.1, d bounds the distance from q to the tangent at p. Observe that $d(p,q) = d(p,r) = 2(1-\alpha)f(\tilde{p})\sin(\frac{\angle pxr}{2})$ and $d = d(p,r) \cdot \sin(\frac{\angle pxr}{2})$. Thus, $d = 2(1-\alpha)f(\tilde{p})\sin^2(\frac{\angle pxr}{2}) = \frac{d(p,q)^2}{2(1-\alpha)f(\tilde{p})}$.

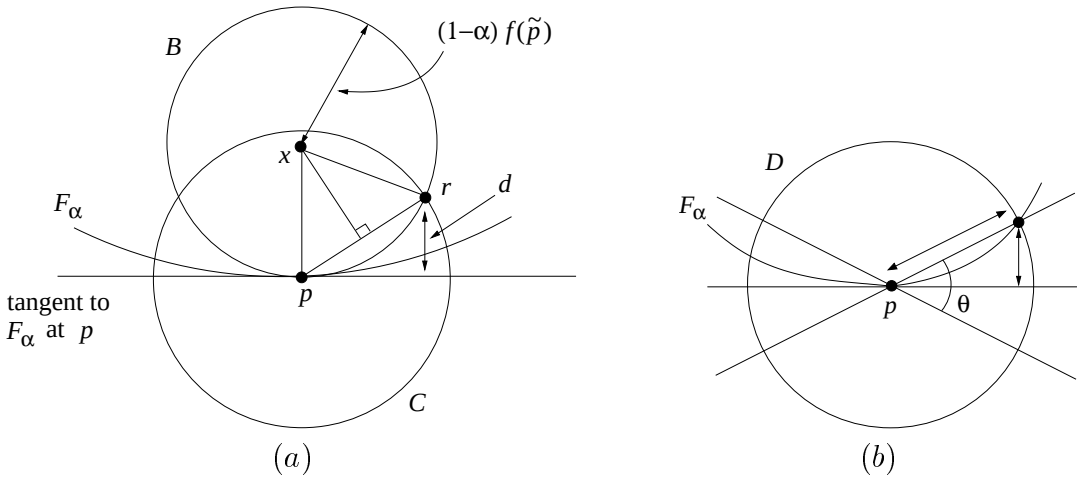

Figure 2.5: Illustration for Lemma 2.2.2.

Consider (ii). Refer to Figure 2.5(b). By (i), the distance between any point in $F_\alpha \cap D$ and the tangent at p is bounded by $\frac{\text{radius}(D)^2}{2(1-\alpha)f(\tilde{p})}$. Let θ be the smallest angle such that $\mathit{cocone}(p,\theta)$ contains $S_\alpha \cap D$. Then $\sin\frac{\theta}{2} \leq \frac{\text{radius}(D)^2}{2(1-\alpha)f(\tilde{p})} \cdot \frac{1}{\text{radius}(D)} = \frac{\text{radius}(D)}{2(1-\alpha)f(\tilde{p})}$. □

Proof of Lemma 2.2.3

Take any point u on $S_\alpha \cap D$. Let ℓ be the tangent to S_α at u. Let ℓ' be the line that is perpendicular to ℓ and passes through u. Let C be the circle centered at p with radius $d(p,u)$. Let A and B be the two tangent circles at p with radius $\frac{(1-\alpha)f(\tilde{p})}{2}$. Let x be the center of A. Without loss of generality, we assume that the tangent to S_α at p is horizontal, A is below B, u lies to the left of p, and the slope of ℓ is positive or infinite. (We ignore the case where the slope of ℓ is zero as there is nothing to prove then.) It follows that the slope of ℓ' is zero or negative. Refer to Figure 2.6.

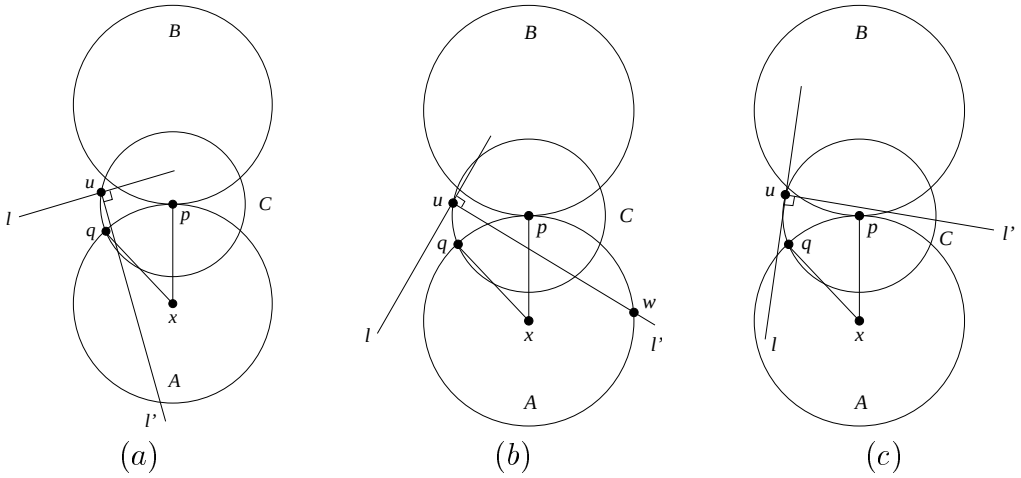

Figure 2.6: Illustration for Lemma 2.2.3.

By Lemma 2.2.1, u lies outside A and B. Let q be the intersection point between C and A on the left of p. Since $d(p,q) = d(p,u) \leq \frac{(1-\alpha)f(\tilde{p})}{4} = \text{radius}(A)/2$, q lies above x. Also, $\angle pxq = 2\sin^{-1}\frac{d(p,u)}{(1-\alpha)f(\tilde{p})}$.

Suppose that ℓ' does not lie above x, see Figure 2.6(a). Since u lies above the support line of qx, the angle between ℓ' and the vertical is less than or equal to $\angle pxq = 2\sin^{-1}\frac{d(p,u)}{(1-\alpha)f(\tilde{p})}$.

Suppose that ℓ' lies above x but not above p, see Figure 2.6(b). We show that this case is impossible. Let w the intersection point between A and ℓ' on the right of p. Note that p lies between u and w and $\angle upw > \pi/2$. If we grow

a disk that lies below l and remains tangent to l at u, the disk will hit S_α at some point different from u when the disk passes through p or earlier. It follows that there is a medial disk M_u of S_α that touches u and lies below l. Observe that the center of M_u lies on the half of ℓ' on the right of u. Furthermore, the center of M_u lies on the line segment uw; otherwise, since $\angle upw > \pi/2$, M_u would contain p, a contradiction. Thus, the distance from $\tilde{p}$ to the center of M_u is less than $\max\{d(p,u), d(p,w)\} + d(p,\tilde{p}) \leq 2 \cdot \text{radius}(A) + \alpha = (1-\alpha)f(\tilde{p}) + \alpha \leq f(\tilde{p})$. However, since the center of M_u is also a point on the medial axis of S, its distance from $\tilde{p}$ should be at least $f(\tilde{p})$, a contradiction.

The remaining case is that ℓ' lies above p, see Figure 2.6(c). Since u lies outside B and the slope of ℓ' is zero or negative, ℓ' lies between p and the center of B. The situation is similar to the previous case where ℓ' lies between p and x. So a similar argument shows that this case is also impossible. ◻

CHAPTER 3

CURVE RECONSTRUCTION

In this chapter, we will present an algorithm to reconstruct polygonal closed curves from noisy samples drawn from a set of smooth closed curves. We assume each input sample point in P follows locally uniform distribution, and the distribution of each sample point is independently identical. We show that the output polygonal reconstruction by our algorithm is faithful to the original curve S with probability approaching to 1 as n increases.

In Section 3.1, we first present our algorithm. Then we states the main theorem of this chapter and give an overview of the analysis of the faithfulness proof in Section 3.2. The details of the analysis appear from Section 3.3 to Section 3.6. Finally we summarize in Section 3.7. Note that the proofs of some technical lemmas are given in the appendix.

3.1 Algorithm

Our algorithm consists of three main steps, Point Estimation, Pruning, and Output. In the Point Estimation step, the algorithm filters out the noise and computes new points that are provably much less noisy than the input samples. Since the sampling density is high, the distances of these new points from S can still be much larger than the distances among them. Thus a direct reconstruction using all of the new points would produce a highly jagged polygonal curve. As a remedy, in the Pruning step, the algorithm decimates the points so that the interpoint distances in the pruned subset is large relative to their distances from S. See Figure 3.1. Finally, in the Output step, we can run any provably good combinatorial curve reconstruction algorithm. We choose to run NN-crust [16]. As it is quite simple, we will then briefly describe its details here. In NN-crust algorithm, each sample s in the given sample points P is connected to two edges. First s is connected to its

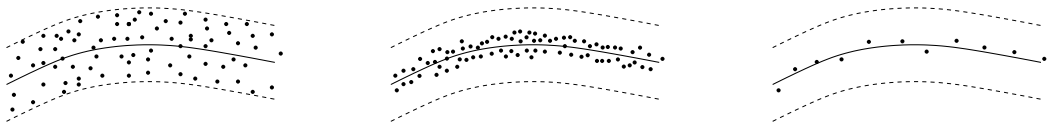

Figure 3.1: The left figure shows the noisy samples. The middle figure shows the new points computed. The right figure shows the remaining points after pruning.

nearest neighbor in P the obtain one edge say e. Then s is connected to the closest sample among all samples u such that su makes an obtuse angle with e.

The following pseudocode gives a high level description of the above three steps and more details of the pruning step. For each point $x \in \mathbb{R}^2$ that does not lie on the medial axis of S, we use $\tilde{x}$ to denote the point on S closest to x. That is, $\tilde{x}$ is the projection of x onto S. (We are not interested in points on the medial axis.)

Point Estimation: For each sample s, we construct a thin rectangle *refined*(s). The long axis of *refined*(s) passes through s and its orientation approximates the normal at $\tilde{s}$. The center of *refined*(s) is the new point s^* desired. The distance $d(s^*, \tilde{s})$ approaches zero as $n \to \infty$.

Pruning: We sort the points s^* in decreasing order of width(*refined*(s)). Then we scan the sorted list and select a subset of center points: when we select a center point s^*, we delete all center points u^* from the sorted list such that $d(s^*, u^*) \leq \text{width}(\textit{refined}(s))^{1/3}$.

Output: We run the NN-crust algorithm on the selected center points and return the output curve.

The main objective of Point Estimation is to align the long axis of *refined*(s) with the normal at $\tilde{s}$. This is instrumental to proving that $d(s^*, \tilde{s})$ approaches zero as $n \to \infty$. The construction of *refined*(s) is done in three steps. We give a highlight first before providing the details.

First, we compute a small disk *initial*(s) centered at s. We prove upper and lower bounds on the radius of *initial*(s), but their ratio is $\Theta(\frac{n^{1/4}}{\ln^{(1+\omega)/4} n})$ which tends to infinity as $n \to \infty$. So *initial*(s) does not provide a reliable estimate of $f(\tilde{s})$.

Second, we grow the disk neighborhood around s until the samples inside the disk fit inside a strip whose width is small relative to the radius of the disk. The final disk is the *coarse neighborhood* of s and it is denoted by $\mathit{coarse}(s)$. The radius of $\mathit{coarse}(s)$ is in the order of $\delta + \text{radius}(\mathit{initial}(s))$. The orientation of the strip approximates the tangent at $\tilde{s}$. Since S can bend quite a lot within $\mathit{coarse}(s)$, the approximation error may be in the order of $\sin^{-1}\delta$. Thus an improved estimate is needed. Third, we shrink $\mathit{coarse}(s)$ to a smaller disk. We take a slab perpendicular to $\mathit{strip}(s)$ bounded by two parallel tangent lines of the shrunken disk. We rotate the slab around s to minimize the spread of the samples inside along the direction of the slab. Because of the minimization of the spread of samples inside, we can show that the orientation of the final slab approximates the normal at $\tilde{s}$ well.

We provide the details of the three steps in POINT ESTIMATION below. Let $\omega > 0$ and $\rho \geq 5$ be two predefined constants.

Initial disk: We compute a disk D centered at s that contains $\ln^{1+\omega} n$ samples. Then we set $\mathit{initial}(s)$ to be the disk centered at s with radius $\sqrt{\text{radius}(D)}$. For sufficiently large n, the radius of D is less than 1, which implies that $\mathit{initial}(s)$ contains D. Figure 3.2 shows an illustration.

Coarse neighborhood: We initialize $\mathit{coarse}(s) = \mathit{initial}(s)$ and compute an infinite strip $\mathit{strip}(s)$ of minimum width that contains all samples inside $\mathit{coarse}(s)$. We grow $\mathit{coarse}(s)$ and maintain $\mathit{strip}(s)$ until $\frac{\text{radius}(\mathit{coarse}(s))}{\text{width}(\mathit{strip}(s))} \geq \rho$. The final disk $\mathit{coarse}(s)$ is the *coarse neighborhood* of s. Figure 3.2 illustrates the growth process.

Refined neighborhood: Let N_s be the upward direction perpendicular to $\mathit{strip}(s)$. The candidate neighborhood $\mathit{candidate}(s, \theta)$ is the slab that contains s in the middle and makes a signed acute angle θ with N_s. The width of $\mathit{candidate}(s, \theta)$ is equal to the minimum of $\sqrt{\text{radius}(\mathit{initial}(s))}$ and $\text{radius}(\mathit{coarse}(s))/3$. The angle θ is positive (resp., negative) if it is on right (resp., left) of N_s. Figure 3.3 shows the initial candidate neighborhood that is perpendicular to $\mathit{strip}(s)$. We enclose the samples in $\mathit{candidate}(s, \theta) \cap$

coarse(s) by two parallel lines that are orthogonal to the direction of *candidate*(s, θ). These two lines form a rectangle *rectangle*(s, θ) with the boundary lines of *candidate*(s, θ). The width of the rectangle *rectangle*(s, θ) is the width of *candidate*(s, θ). The height of *rectangle*(s, θ) is its length along the direction of *candidate*(s, θ). We vary θ within the range $[-\pi/10, \pi/10]$ to find an orientation that minimizes the height of *rectangle*(s, θ). Figure 3.3 illustrates the rotation and the bounding rectangle. Let θ^* be the minimizing angle. The *refined neighborhood* of s is *rectangle*(s, θ^*) and is denoted by *refined*(s). We return the center point s^* of *refined*(s).

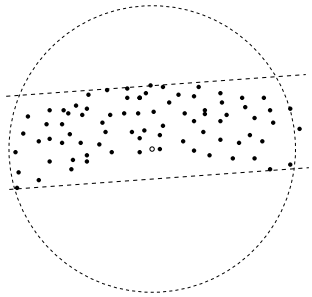

Figure 3.2: On the left, the white dot is the sample s, the inner disk is D, and the outer disk is *initial*(s). On the right, we grow *initial*(s) until *strip*(s) has a relatively large aspect ratio. The final disk is *coarse*(s).

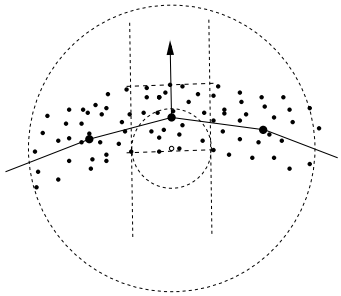

Figure 3.3: On the left, the initial candidate neighborhood is the one perpendicular to *strip*(s). On the right, as we rotate the candidate neighborhood, we maintain the smallest bounding rectangle of all samples inside.

A few remarks are in order. Recall that $\min_{x \in S} f(x)$ is assumed to be 1. For sufficiently large n (i.e., when the sampling is dense enough), the radius of *initial*(s) is less than 1. So in the REFINED NEIGHBORHOOD step, $\sqrt{\text{radius}(\textit{initial}(s))} >$

radius($initial(s)$). Clearly, $coarse(s)$ contains $initial(s)$. So both of the widths of $candidate(s, \theta)$ and $refined(s)$ are at most $\sqrt{\text{radius}(initial(s))} < 1$ and at least radius($initial(s)$)/3.

3.2 Overview of analysis

Our goal is to prove the following result:

Main Theorem *Assume that $\delta \leq 1/(25\rho^2)$ and $\rho \geq 5$. Let n be the number of noisy samples from a smooth closed curve. For sufficiently large n, our algorithm computes a polygonal closed curve that has the following properties with probability at least $1 - O(n^{-\Omega(\frac{\ln^{\omega} n}{f_{\max}} - 1)})$.*

- *For each output vertex s^*, $d(s^*, \tilde{s}) = O((\frac{\ln^{1+\omega} n}{n})^{1/8} f(\tilde{s})^{1/4})$.*
- *For each output edge r^*s^*, the angle between r^*s^* and the tangent at $\tilde{s}$ is $O((\frac{\ln^{1+\omega} n}{n})^{1/48} f(\tilde{s})^{25/24})$.*
- *The output curve is homeomorphic to the smooth closed curve.*

We first give an overview of the proof strategies here before diving into details later. The hardest part is to argue that the point s^* that we estimate for the sample s indeed lies very closely to the curve. To illustrate the intuition, we assume that the curve is a flat horizontal segment locally at $\tilde{s}$. See Figure 3.4(a). So the noisy samples in the local neighborhood lie within a band B of width 2δ. Thus the final $coarse(s)$ must have radius $\Theta(\rho\delta + \text{radius}(initial(s)))$ in order to meet the stopping criterion of growing $coarse(s)$. Next, we would like to argue that the slope of $strip(s)$ approximates the slope of the tangent at $\tilde{s}$. We prove this by contradiction and assume that $strip(s)$ is tilted a lot. So a significant area of B lies outside $strip(s)$ as shown in Figure 3.4(b). Our goal is to show that this area contains a noisy sample with high probability. Therefore, with high probability, $strip(s)$ cannot be much tilted from the horizontal. The details are in Section 3.3.

Directly discussing the emptiness of an arbitrary area (whether it contains a noisy sample or not) is quite hard given the continuous distributions. We get around this

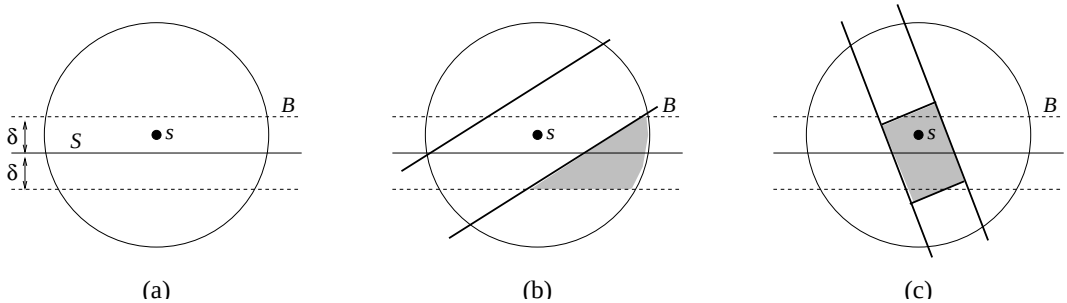

Figure 3.4: The left figure shows $coarse(s)$, the noise band B, and S. In the middle figure, the bold strip is $strip(s)$ and the shaded area is the significant area of B outside $strip(s)$. The shaded area should be non-empty with high probability. In the right figure, the shaded rectangle is the candidate rectangle.

by decomposing the space around S into small cells by making use of the properties in Chapter 2. Since the cells have more regular shape, we can show that each cell is non-empty with high probability and we can also bound the diameters of the cells. The cell diameter approaches zero as the sampling density increases. The bound on the cell diameter enables us to show that the area of B outside $strip(s)$ in Figure 3.4(b) contains a cell. So the area contains a noisy sample with high probability.

The next step is to construct the refined neighborhood of s so as to obtain an improved estimate of the normal at $\tilde{s}$. This is done by rotating a candidate rectangle to minimize its height. See Figure 3.4(c). The width of the candidate rectangle is set to be the minimum of $\sqrt{\text{radius}(initial(s))}$ and $\text{radius}(coarse(s))/3$. Clearly, we want the width to be small in order to generate a large variation in the height even when we have a small angular deviation from the normal at $\tilde{s}$. In fact, we want to show that $\text{radius}(initial(s))$ approaches zero as the sampling density increases. Recall that $initial(s)$ is generated by identifying the $\ln^{1+\omega} n$ nearest samples to s. We are to show that the number of samples inside a cell is at least $\ln^{1+\omega} n$ with high probability. Thus $\text{radius}(initial(s))$ is no more than the cell diameter. In Figure 3.4(c), when we rotate the candidate rectangle, its upper and lower sides may invade the interior of the band B. This is because there may not be any noisy sample on the band boundary. Still, we want to keep the upper and lower sides of

the candidate rectangle near the band boundary, otherwise we would not have a big increase in height despite the angular deviation from the normal at $\tilde{s}$. Fortunately, as the cells are non-empty with high probability, the gaps between the upper and lower sides and the band boundary must be too narrow for a single cell to fit in. The details are in Section 3.4.

We have not discussed one important phenomenon so far. Since δ is unknown, it may be arbitrarily small. In this case, radius($\mathit{coarse}(s)$) is only lower bounded by radius($\mathit{initial}(s)$) as we grow $\mathit{coarse}(s)$ from $\mathit{initial}(s)$. Thus we need to establish a lower bound on radius($\mathit{initial}(s)$), and hence radius($\mathit{coarse}(s)$). We construct another decomposition of the space around S into slabs. Then by upper bounding the number of samples in each slab, we can lower bound radius($\mathit{initial}(s)$) by the slab "width". The corresponding details are in the following section, in which the radii of $\mathit{initial}(s)$ and $\mathit{coarse}(s)$ for each sample s will be bounded from above and below.

In Section 3.5, we obtain the homeomorphism result by extending the NN-crust analysis. In Section 3.6, we put everything together to prove the main theorem of this chapter.

3.3 Coarse neighborhood

In this section, we bound the radii of $\mathit{initial}(s)$ and $\mathit{coarse}(s)$ for each sample s. Then we show that $\mathit{strip}(s)$ provides a rough estimate of the slope of the tangent to S at $\tilde{s}$. Recall that $\lambda_k = \sqrt{\frac{k^2 \ln^{1+\omega} n}{n}}$.

3.3.1 Radius of $\mathit{initial}(s)$

Lemma 3.3.1 *Let h be a constant less than $\sqrt{\frac{1}{3\kappa_1}}$ and let m be a constant greater than $\sqrt{\frac{2}{\kappa_2}}$, where κ_1 and κ_2 are the constants in Lemma 2.3.4. Let $\psi_h = \lambda_h/3$ and $\psi_m = \sqrt{14\lambda_m}$. Let s be a sample. If $\delta \leq 1/8$, $\lambda_h \leq 1/12$, and $\lambda_m \leq 1/12$, then*

$$\psi_h \sqrt{f(\tilde{s})} \leq \text{radius}(\mathit{initial}(s)) \leq \psi_m \sqrt{f(\tilde{s})}$$

with probability at least $1 - O(n^{-\Omega(\ln^{\omega} n)})$.

Proof. Let D be the disk centered at s that contains $\ln^{1+\omega}$ samples. We first prove the upper bound. Take a λ_m-grid such that s lies on the normal segment at the cut-point c_0. Let C be the λ_m-cell between the normal segments at c_0 and c_1 that contains s. By Lemma 2.3.6(iii), C contains at least $2\ln^{1+\omega} n$ samples with probability at least $1-n^{-\Omega(\ln^{\omega} n)}$. Since D contains $\ln^{1+\omega} n$ samples, radius(D) is less than the diameter of C with probability at least $1 - n^{-\Omega(\ln^{\omega} n)}$. By Lemma 2.3.2, radius(D) $\leq 14\lambda_m f(c_0) = 14\lambda_m f(\tilde{s})$. Hence radius($\mathit{initial}(s)$) $= \sqrt{\text{radius}(D)} \leq \sqrt{14\lambda_m f(\tilde{s})}$.

Next, we prove the lower bound. Take a λ_h-partition such that s lies on the normal segment at the cut-point c_0. Consider the cut-points c_j for $-1 \leq j \leq 1$. (We use c_{-1} to denote the last cut-point picked.) We have $d(c_{-1}, c_0) \leq |S(c_{-1}, c_0)| \leq 3\lambda_h^2 f(c_{-1}) < 0.03 f(c_{-1})$ as $\lambda_h \leq 1/12$. The Lipschitz condition implies that

$$f(c_{-1}) \geq f(c_0)/1.03 > 0.8 f(c_0). \tag{3.1}$$

Let d_{-1} and d_1 be the distances from s to the support lines of the normal segments at c_{-1} and c_1, respectively. By Lemma 2.3.3,

$$d_{-1} \geq \frac{|S(c_{-1}, c_0)|}{6} \geq \frac{\lambda_h^2 f(c_{-1})}{6} \overset{(3.1)}{>} \frac{\lambda_h^2 f(c_0)}{8},$$

$$d_1 \geq \frac{|S(c_0, c_1)|}{6} \geq \frac{\lambda_h^2 f(c_0)}{6}.$$

By Lemma 2.3.6(ii), the λ_h-slabs between c_{-1} and c_0 and between c_0 and c_1 contain at most $\ln^{1+\omega} n/3$ points with probability at least $1 - O(n^{-\Omega(\ln^{\omega} n)})$. Hence, for D to contain $\ln^{1+\omega} n$ points, radius(D) $> \max\{d_{-1}, d_1\} \geq \lambda_h^2 f(c_0)/6$. Note that $f(\tilde{s}) = f(c_0)$ as $\tilde{s} = c_0$ by construction. It follows that radius($\mathit{initial}(s)$) $= \sqrt{\text{radius}(D)} > \lambda_h \sqrt{f(\tilde{s})}/3$. ∎

3.3.2 Radius of $\mathit{coarse}(s)$

In this section, we prove an upper bound and a lower bound on the radius of $\mathit{coarse}(s)$.

Lemma 3.3.2 *Assume $\rho \geq 4$ and $\delta \leq 1/(25\rho^2)$. Let m be the constant and ψ_m be the parameter in Lemma 3.3.1. Let s be a sample. If $\lambda_m \leq 1/(504\rho^2)$, then*

$$\text{radius}(\mathit{coarse}(s)) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$$

with probability at least $1 - O(n^{-\Omega(\ln^\omega n)})$.

Proof. Let s_1 and s_2 be points on S_δ^+ and S_δ^- such that $\tilde{s_1} = \tilde{s_2} = \tilde{s}$. Let D be the disk centered at s with radius $5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$. By Lemma 3.3.1, $\psi_m\sqrt{f(\tilde{s})} \geq \text{radius}(\mathit{initial}(s))$, so D contains $\mathit{initial}(s)$ with probability at least $1 - O(n^{\Omega(\ln^\omega n)})$. We are to show that $\mathit{coarse}(s)$ cannot grow beyond D. First, since $\lambda_m \leq 1/(504\rho^2)$,

$$\psi_m = \sqrt{14\lambda_m} \leq 1/(6\rho) \leq 1/24.$$

Observe that both s_1 and s_2 lie inside D. Since $5\rho\delta \leq 1/(5\rho) \leq 1/20$ and $\psi_m \leq 1/24$, $\text{radius}(D) < (1-\delta)f(\tilde{s})$. Thus, the distance between any two points in $D \cap S_\delta^+$ is less than $2(1-\delta)f(\tilde{s})$. By Lemma 2.2.2(i), the maximum distance between $D \cap S_\delta^+$ and the tangent to S_δ^+ at s_1 is at most $\frac{(5\rho\delta+\psi_m\sqrt{f(\tilde{s})})^2}{2(1-\delta)f(\tilde{s})} \leq \frac{(5\rho\delta\sqrt{f(\tilde{s})}+\psi_m\sqrt{f(\tilde{s})})^2}{2(1-\delta)f(\tilde{s})}$ as $f(\tilde{s}) \geq 1$. Thus, this distance is upper bounded by $\frac{(5\rho\delta+\psi_m)^2}{2(1-\delta)}$ which is less than $0.51(5\rho\delta + \psi_m)^2$ as $\delta \leq 1/(25\rho^2)$. The same is also true for $D \cap S_\delta^-$. It follows that the samples inside D lie inside a strip of width at most $2\delta + 1.1(5\rho\delta + \psi_m)^2 = 2\delta + 1.1(5\rho)^2\delta^2 + 2.2(5\rho)\psi_m\delta + 1.1\psi_m^2$. Since $\delta \leq 1/(25\rho^2)$ and $\psi_m \leq 1/(6\rho)$, we have $1.1(5\rho)^2\delta^2 \leq 1.1\delta$, $2.2(5\rho)\psi_m\delta < 1.84\delta$, and $1.1\psi_m^2 < \psi_m/\rho$. We conclude that the strip width is no more than $2\delta + 1.1\delta + 1.84\delta + \psi_m/\rho < 5\delta + \psi_m/\rho \leq \text{radius}(D)/\rho$. This shows that $\mathit{coarse}(s)$ cannot grow beyond D. ◻

Next, we bound $\text{radius}(\mathit{coarse}(s))$ from below. We use $f_{\max}$ to denote $\max_{x \in S} f(x)$.

Lemma 3.3.3 *Assume that $\delta \leq 1/8$ and $\rho \geq 4$. Let h be the constant in Lemma 3.3.1. Let s be a sample. If $\lambda_h \leq 1/32$, then*

$$\text{radius}(\mathit{coarse}(s)) \geq \max\{2\sqrt{\rho}\delta, \text{radius}(\mathit{initial}(s))\}$$

with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.

Proof. Since *coarse*(s) is grown from *initial*(s), radius(*coarse*(s)) $\geq$ radius(*initial*(s)). We are to prove that radius(*coarse*(s)) $\geq 2\sqrt{\rho}\delta$. Let D be the disk that has center s and radius radius(*coarse*(s))$/\sqrt{\rho}$. Let X be the disk centered at $\tilde{s}$ with radius δ. Note that $s \in X$ and X is tangent to S_δ^+ and S_δ^-. Since $\delta \leq 1/8$ and $f(\tilde{s}) \geq 1$, $f(\tilde{s}) - \delta > \delta$ and so Lemma 2.2.1 implies that X lies inside the finite region bounded by S_δ^+ and S_δ^-.

Suppose that radius(*coarse*(s)) $< 2\sqrt{\rho}\delta$. Then radius$(D) < 2\delta$. If D contains X, X is a disk inside $D \cap X$ with radius at least radius$(D)/2$. If D does not contain X, then since $s \in X$, $D \cap X$ contains a disk with radius radius$(D)/2$. The width of *strip*(s) is less than or equal to radius(*coarse*(s))$/\rho$ = radius$(D)/\sqrt{\rho}$. Thus, $(D \cap X) -$ *strip*(s) contains a disk Y such that

$$\text{radius}(Y) \geq (\frac{1}{4} - \frac{1}{4\sqrt{\rho}}) \cdot \text{radius}(D) \geq \frac{\text{radius}(D)}{8}.$$

Note that Y is empty and Y lies inside the finite region bounded by S_δ^+ and S_δ^-. Take a point $p \in Y$. Since $p \in Y \subseteq D$ and radius$(D) < 2\delta$, $d(\tilde{p}, \tilde{s}) \leq d(p, \tilde{p}) + d(s, \tilde{s}) + d(p, s) \leq 4\delta \leq 1/2$ as $\delta \leq 1/8$. The Lipschitz condition implies that $f(\tilde{p}) \leq 3f(\tilde{s})/2$. Observe that radius$(D)$ = radius(*coarse*(s))$/\sqrt{\rho} \geq$ radius(*initial*(s))$/\sqrt{\rho}$. Thus, Lemma 3.3.1 implies that radius$(Y) \geq$ radius$(D)/8 \geq \lambda_h\sqrt{f(\tilde{s})}/(24\sqrt{\rho}) > \lambda_h\sqrt{f(\tilde{p})}/(30\sqrt{\rho})$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n)})$. Let $\beta = \lambda_h/(420\sqrt{\rho f_{\max}})$. Then radius$(Y) > 14\beta f(\tilde{p})$. By Lemma 2.3.2, Y contains a β-cell. By Lemma 2.3.6(i), this β-cell is empty with probability at most $n^{-\Omega(\ln^\omega n/f_{\max})}$. This implies that radius(*coarse*(s)) $< 2\sqrt{\rho}\delta$ occurs with probability at most $O(n^{-\Omega(\ln^\omega n/f_{\max})})$. □

3.3.3 Rough tangent estimate: *strip*(s)

In this section, we prove that the slope of *strip*(s) is a rough estimate of the slope of the tangent at $\tilde{s}$. We need the following technical lemma about various properties of *coarse*(s) and S_α inside *coarse*(s). Its proof can be found in the appendix.

Lemma 3.3.4 *Assume $\rho \geq 5$ and $\delta \leq 1/(25\rho^2)$. Let m be the constant and ψ_m be the parameter in Lemma 3.3.1. Let s be a sample. If $2\sqrt{\rho}\delta \leq \text{radius}(\textit{coarse}(s)) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$ and $\psi_m \leq 1/100$, then for any S_α and for any point $x \in S_\alpha \cap \textit{coarse}(s)$, the following hold:*

(i) $5\rho\delta + \psi_m \leq 0.05$, $\frac{5\rho\delta+\psi_m}{2(1-\delta)} \leq 0.03$, *and* $\frac{5\rho\delta+\psi_m+2\delta}{2(1-\delta)} \leq 0.03$,

(ii) $S_\alpha \cap \textit{coarse}(s)$ *consists of one connected component,*

(iii) the angle between the normals at s and x is at most $2\sin^{-1}\frac{5\rho\delta+\psi_m+2\delta}{(1-\delta)} \leq 2\sin^{-1}(0.06)$,

(iv) $x \in \textit{cocone}(s_1, 2\sin^{-1}\frac{5\rho\delta+\psi_m+2\delta}{2(1-\delta)}) \subseteq \textit{cocone}(s_1, 2\sin^{-1}(0.03))$ *where s_1 is the point on S_α such that $\tilde{s_1} = \tilde{s}$.*

(v) $0.9f(\tilde{s}) < f(\tilde{x}) < 1.1f(\tilde{s})$,

(vi) if x lies on the boundary of $\textit{coarse}(s)$, the distance between s and the orthogonal projection of x onto the tangent at s is at least $0.8 \cdot \text{radius}(\textit{coarse}(s))$, *and*

(vii) for any $y \in S_\alpha \cap \textit{coarse}(s)$, the acute angle between xy and the tangent at x is at most $\sin^{-1}(6\rho\delta + 1.2\psi_m)) \leq \sin^{-1}(0.06)$.

We highlight the key ideas before giving the proof of Lemma 3.3.5. Let $\mathcal{B}$ be the region between S_δ^+ and S_δ^- inside $\textit{coarse}(s)$. If $\textit{strip}(s)$ makes a large angle with the tangent at $\tilde{s}$, $\textit{strip}(s)$ would cut through $\mathcal{B}$ in the middle. In this case, if $\mathcal{B} \cap \textit{strip}(s)$ is narrow, there would be a lot of areas in $\mathcal{B}$ outside $\textit{strip}(s)$. But these areas must be empty. Such areas occur with low probability. Otherwise, if $\mathcal{B} \cap \textit{strip}(s)$ is wide, we show that $\textit{strip}(s)$ can be rotated to reduce its width further, a contradiction. We give the detailed proof below.

Lemma 3.3.5 *Assume that $\rho \geq 5$ and $\delta \leq 1/(25\rho^2)$. Let m be the constant and ψ_m be the parameter in Lemma 3.3.1. Let s be a sample. For sufficiently large n, the acute angle between the tangent at $\tilde{s}$ and the direction of $\textit{strip}(s)$ is at most* $3\sin^{-1}\frac{5\rho\delta+\psi_m+2\delta}{(1-\delta)} + \sin^{-1}(6\rho\delta + 1.2\psi_m) \leq 4\sin^{-1}(0.06)$ *with probability at least* $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.

Proof. Let ℓ_1 and ℓ_2 be the lower and upper bounding lines of $strip(s)$. Without loss of generality, we assume that the normal at $\tilde{s}$ is vertical, the slope of $strip(s)$ is non-negative, $S_\delta^- \cap coarse(s)$ lies below $S_\delta^+ \cap coarse(s)$, and $\psi_m \leq 1/100$ for sufficiently large n. Let h and m be the constants and ψ_h and ψ_m be the parameters in Lemma 3.3.1. We first assume that $\max\{2\sqrt{\rho}\delta, \psi_h\sqrt{f(\tilde{s})}\} \leq \text{radius}(coarse(s)) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$ and take the probability of its occurrence into consideration later. As a short hand, we use η_1 to denote $\frac{5\rho\delta+\psi_m+2\delta}{(1-\delta)}$ and η_2 to denote $6\rho\delta + 1.2\psi_m$.

Observe that both ℓ_1 and ℓ_2 must intersect the space that lies between S_δ^+ and S_δ^- inside $coarse(s)$. Otherwise, we can squeeze $strip(s)$ and reduce its width, a contradiction. If ℓ_1 intersects $S_\alpha \cap coarse(s)$ twice for some α, then ℓ_1 is parallel to the tangent at some point on $S_\alpha \cap coarse(s)$. By Lemma 3.3.4(iii), the direction of $strip(s)$ makes an angle at most $2\sin^{-1}\eta_1$ with the horizontal and we are done. Similarly, we are done if ℓ_2 intersects $S_\alpha \cap coarse(s)$ twice for some α. The remaining case is that both ℓ_1 and ℓ_2 intersect $S_\alpha \cap coarse(s)$ for any α at most once. Suppose that the acute angle between the direction of $strip(s)$ and the horizontal is more than $3\sin^{-1}\eta_1 + \sin^{-1}\eta_2$. We show that this occurs with probability $O(n^{-\Omega(\ln^\omega n/f_{\max})})$.

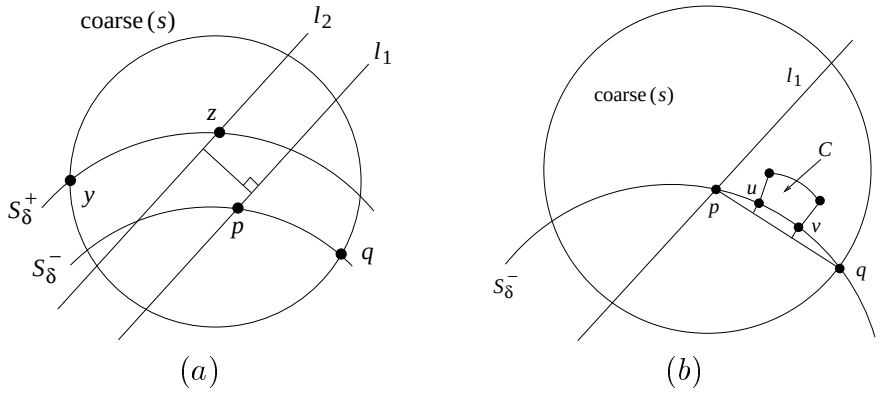

Figure 3.5: Figure (a) illustrates that $S_\delta^-(p,q)$ lies below ℓ_1. Figure (b) illustrates our choice of a cell C that lies below ℓ_1.

Let q be the right intersection point between S_δ^- and the boundary of $coarse(s)$. If ℓ_1 intersects $S_\delta^- \cap coarse(s)$, let p denote the intersection point; otherwise, let p denote the leftmost intersection point between S_δ^- and the boundary of $coarse(s)$. Refer to Sigure 3.5(a). We claim that $S_\delta^-(p,q)$ lies below ℓ_1. If ℓ_1 does not intersect

$S_\delta^- \cap \mathit{coarse}(s)$, then this is clearly true. Otherwise, by Lemma 3.3.4(iii), the magnitude of the slope of the tangent at p is at most $2\sin^{-1}\eta_1$. Since the slope of ℓ_1 is more than $3\sin^{-1}\eta_1 + \sin^{-1}\eta_2$, S_δ^- crosses ℓ_1 at p from above to below. So $S_\delta^-(p,q)$ lies below ℓ_1.

We show that $d(p,q) \leq \psi_h\sqrt{f(\tilde{s})}/2$ with probability at least $1 - n^{-\Omega(\ln^\omega n/f_{\max})}$. Notice that pq is parallel to the tangent to S_δ^- at some point on $S_\delta^-(p,q)$. By Lemma 3.3.4(iii), the tangent to $S_\delta^-(p,q)$ turns by an angle at most $4\sin^{-1}(0.06) < \pi/2$ from p to q. This implies that $S_\delta^-(p,q)$ is monotone with respect to the direction perpendicular to pq. We divide pq into three equal segments. Let u and v be the intersection points between $S_\delta^-(p,q)$ and the perpendiculars of pq at the dividing points. Assume that v follows u along $S_\delta^-(p,q)$. Refer to Figure 3.5(b). Suppose that $d(p,q) > \psi_h\sqrt{f(\tilde{s})}/2$. Then

$$|S_\delta^-(u,v)| \geq \frac{d(p,q)}{3} \geq \frac{\psi_h\sqrt{f(\tilde{s})}}{6}. \tag{3.2}$$

Since $f(\tilde{u}) < 1.1f(\tilde{s})$ by Lemma 3.3.4(v), $|S_\delta^-(u,v)| > \psi_h\sqrt{f(\tilde{u})}/7$. Consider a $(\lambda_k/\sqrt{f_{\max}})$-grid where $k = h/294$ and $\tilde{u}$ is a cut-point. (Note that $\lambda_k = \psi_h/98$.) Let C be the $(\lambda_k/\sqrt{f_{\max}})$-cell that touches $S_\delta^-(u,v)$ and the normal segment through u. By Lemma 2.3.2, the diameter of C is at most $14\lambda_k\sqrt{f(\tilde{u})} = \psi_h\sqrt{f(\tilde{u})}/7 < |S_\delta^-(u,v)|$. So the bottom side of C lies within $S_\delta^-(u,v)$. Let $\mathcal{R}$ be the region inside $\mathit{coarse}(s)$ that lies below ℓ_1 and above $S_\delta^-(p,q)$. From any point $x \in S_\delta^-(u,v) \cap C$, if we shoot a ray along the normal at x into $\mathcal{R}$, either the ray will leave C first or the ray will hit ℓ_1 or the boundary of $\mathit{coarse}(s)$ in $\mathcal{R}$. We are to prove that the distances from x to ℓ_1 and the boundary of $\mathit{coarse}(s)$ in $\mathcal{R}$ are more than $2\lambda_k\delta \geq 2\lambda_k\delta/\sqrt{f_{\max}}$. This shows that the ray always leaves C first, so C lies completely inside $\mathcal{R}$. Then the upper bound on $d(p,q)$ follows as C is empty with probability at most $n^{-\Omega(\ln^\omega n/f_{\max})}$ by Lemma 2.3.6(i).

Consider the distance from any point $x \in S_\delta^-(u,v)$ to ℓ_1. By Lemma 3.3.4(iii), the angle between ℓ_1 and the tangent at p (measured by rotating ℓ_1 in the clockwise direction) is at least $3\sin^{-1}\eta_1 + \sin^{-1}\eta_2 - 2\sin^{-1}\eta_1 = \sin^{-1}\eta_1 + \sin^{-1}\eta_2$ and at most $\pi/2 + 2\sin^{-1}\eta_1$. By Lemma 3.3.4(vii), the acute angle between px and the tangent at p is at most $\sin^{-1}\eta_2$. So the angle between px and ℓ_1 is at least $\sin^{-1}\eta_1$

and at most $\pi/2 + 2\sin^{-1}\eta_1 + \sin^{-1}\eta_2$. This implies that the distance from x to ℓ_1 is at least $d(p,x) \cdot \min\{\eta_1\,,\, \cos(2\sin^{-1}\eta_1 + \sin^{-1}\eta_2)\}$. By Lemma 3.3.4(i), $\eta_1 \leq 0.06 < \cos(3\sin^{-1}(0.06)) \leq \cos(2\sin^{-1}\eta_1 + \sin^{-1}\eta_2)$. Therefore, the distance from x to ℓ_1 is at least $d(p,x) \cdot \eta_1 > 5\rho\delta \cdot d(p,x) \geq 25\delta \cdot (d(p,q)/3) \overset{(3.2)}{>} 4\delta\psi_h\sqrt{f(\tilde{s})}$. Since $\lambda_k = \psi_h/98$, this distance is greater than $2\lambda_k\delta$.

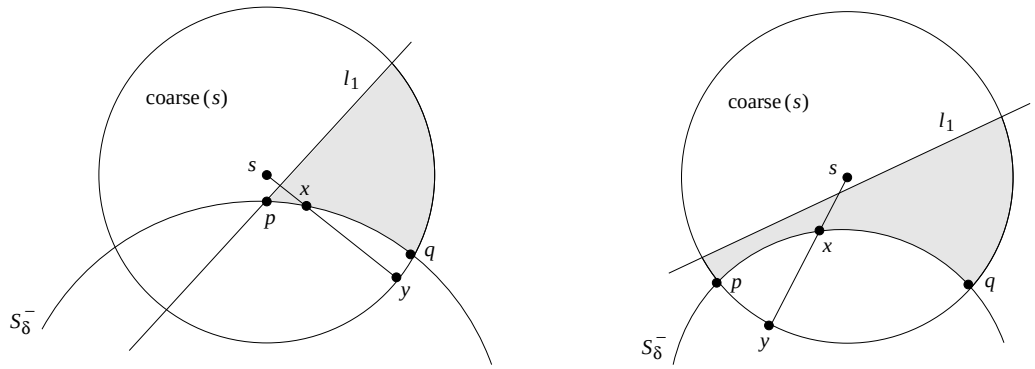

Figure 3.6: The shaded region denotes $\mathcal{R}$ in both figures. In figure (a), q is the closest point in R to x. In figure (b), p or q is the closet point in R to x.

Next, we consider the distance d from any point $x \in S_\delta^-(u,v)$ to the boundary of $coarse(s)$ in $\mathcal{R}$. Take a radius sy of $coarse(s)$ that passes through x. Suppose that y lies outside $\mathcal{R}$. Refer to Figure 3.6. If ℓ_1 intersects $S_\delta^- \cap coarse(s)$ at p (Figure 3.6(a)), then $d = d(q,x)$. If ℓ_1 does not intersect $S_\delta^- \cap coarse(s)$ (Figure 3.6(b)), then $d = \min\{d(p,x)\,,\, d(q,x)\}$. Thus, by (3.2), $d \geq d(p,q)/3 \geq \psi_h\sqrt{f(\tilde{s})}/6 > 2\lambda_k\delta$. The remaining possibility is that y lies on the boundary of $\mathcal{R}$. Then either sy is tangent to S_δ^- at x or sy intersects $S_\delta^- \cap coarse(s)$ at least twice. So xy is parallel to the tangent at some point on $S_\delta^- \cap coarse(s)$. By Lemma 3.3.4(iii), the acute angle between xy and the tangent at x is at most $4\sin^{-1}\eta_1$. By Lemma 3.3.4(vii), the acute angle between qx and the tangent at x is at most $\sin^{-1}\eta_2$. So the angle between qx and xy is at most $4\sin^{-1}\eta_1 + \sin^{-1}\eta_2$. It follows that $d = d(x,y) \geq d(q,x) \cdot \cos(4\sin^{-1}\eta_1 + \sin^{-1}\eta_2) \geq d(q,x) \cdot \cos(5\sin^{-1}(0.06)) > 0.9 \cdot d(q,x) \geq 0.9 \cdot (d(p,q)/3) \geq 0.15\psi_h\sqrt{f(\tilde{s})} > 2\lambda_k\delta$.

In all, C lies inside $\mathcal{R}$. So C must be empty which occurs with probability at most $n^{-\Omega(\ln^\omega n/f_{\max})}$ by Lemma 2.3.6(i). It follows that $d(p,q) \leq \psi_h\sqrt{f(\tilde{s})}/2$ with probability at least $1 - n^{-\Omega(\ln^\omega n/f_{\max})}$. By Lemma 3.3.4(vi), the horizontal distance between q and the left intersection point between S_δ^- and the boundary of $coarse(s)$ is at least $1.6 \cdot \text{radius}(coarse(s)) \geq 1.6\psi_h\sqrt{f(\tilde{s})} > d(p,q)$. We conclude that ℓ_1 intersects $S_\delta^- \cap coarse(s)$ exactly once at p.

Refer to Figure 3.7. Let y be the leftmost intersection point between S_δ^+ and the boundary of $coarse(s)$. Symmetrically, we can also show that ℓ_2 intersects $S_\delta^+ \cap coarse(s)$ exactly once at some point z, $S_\delta^+(y,z)$ lies above ℓ_2, and $d(y,z) \leq \psi_h\sqrt{f(\tilde{s})}/2$ with probability at least $1 - n^{-\Omega(\ln^\omega n/f_{\max})}$.

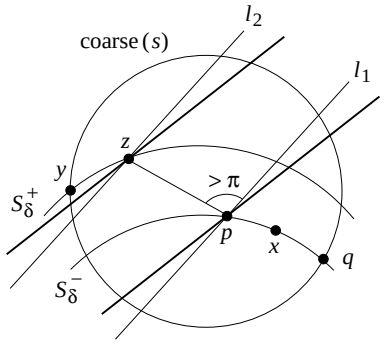

Figure 3.7: Rotating ℓ_1 and ℓ_2 slightly in the clockwise direction decreases the width of $strip(s)$.

Consider the projections of $S_\delta^+(y,z)$ and $S_\delta^-(p,q)$ onto the horizontal diameter of $coarse(s)$ through s. By Lemma 3.3.4(vi), the projections of y and q are at distance at least $0.8 \cdot \text{radius}(coarse(s))$ from s. Thus, the distance between the projections of $S_\delta^+(y,z)$ and $S_\delta^-(p,q)$ is at least $1.6 \cdot \text{radius}(coarse(s)) - d(p,q) - d(y,z) \geq 1.6 \cdot \text{radius}(coarse(s)) - \psi_h\sqrt{f(\tilde{s})} \geq 1.6 \cdot \text{radius}(coarse(s)) - \text{radius}(coarse(s)) > \text{radius}(coarse(s))/\rho$. That is, this distance is greater than the width of $strip(s)$. But then we can rotate ℓ_1 and ℓ_2 around p and z, respectively, in the clockwise direction to reduce the width of $strip(s)$ while not losing any sample inside $coarse(s)$. See Figure 3.7. This is impossible. It follows that, under the condition that $\max\{2\sqrt{\rho}\delta, \psi_h\sqrt{f(\tilde{s})}\} \leq \text{radius}(coarse(s)) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$, the acute angle between the direction of $strip(s)$ and the tangent at $\tilde{s}$ is at most $3\sin^{-1}\eta_1 + \sin^{-1}\eta_2$ with probability at least $1 - O(n^{\Omega(\ln^\omega n/f_{\max})})$. By Lemmas 3.3.1, 3.3.2, and 3.3.3,

the inequalities $\max\{2\sqrt{\rho}\delta, \psi_h\sqrt{f(\tilde{s})}\} \leq \text{radius}(\mathit{coarse}(s)) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$ hold with probability at least $1 - O(n^{\Omega(\ln^{\omega} n)/f_{\max}})$. So the lemma follows. □

3.4 Refined neighborhood

The results in Section 3.3 show that after the step COARSE NEIGHBORHOOD, the algorithm already has a normal estimate at each noisy sample with an error in the order of $\delta + \psi_m$. However, this error bound does not tend to zero as the sampling density increases. This explains the need for the step REFINED NEIGHBORHOOD in the algorithm. This step will improve the normal estimate so that the error tends to zero as the sampling density increases. This will allow us to prove the pointwise convergence.

We introduce some notations. In the step REFINED NEIGHBORHOOD, we align $\mathit{candidate}(s, \theta)$ with the normal at $\tilde{s}$ by varying θ within $[-\pi/10, \pi/10]$. Recall that θ is the signed acute angle between the upward direction of $\mathit{candidate}(s, \theta)$ and N_s, where N_s is the upward direction perpendicular to $\mathit{strip}(s)$. Let $\mathit{angle}(\mathit{strip}(s))$ denote the signed acute angle between N_s and the upward normal at $\tilde{s}$. If N_s points to the right of the upward normal at $\tilde{s}$, $\mathit{angle}(\mathit{strip}(s))$ is positive. Otherwise, $\mathit{angle}(\mathit{strip}(s))$ is negative. We define $\theta_s = \theta + \mathit{angle}(\mathit{strip}(s))$. That is, θ_s is the signed acute angle between the upward direction of $\mathit{candidate}(s, \theta)$ and the upward normal at $\tilde{s}$. The sign of θ_s is determined in the same way as $\mathit{angle}(\mathit{strip}(s))$. For any S_α and for any point $p \in S_\alpha \cap \mathit{candidate}(s, \theta)$, let γ_p be the signed acute angle between the upward direction of $\mathit{candidate}(s, \theta)$ and the upward normal at $\tilde{p}$. The sign of γ_p is determined in the same way as $\mathit{angle}(\mathit{strip}(s))$.

We need the following two technical lemmas. Their proofs can be found in the appendix. There are two main results in Lemma 3.4.1. First, we show that the range of rotation $[-\pi/10, \pi/10]$ of $\mathit{candidate}(s, \theta)$ covers the normal direction at $\tilde{s}$. Second, we relate γ_p to θ_s. This is useful because we will see that for a proper choice of p, the height of $\mathit{candidate}(s, \theta)$ is directly related to γ_p (and hence to θ_s). We will need to focus on a smaller area inside $\mathit{candidate}(s, \theta)$. Lemma 3.4.2 bounds distances and angles involving points on S_α inside this smaller area.

Lemma 3.4.1 *Assume that $\delta \leq 1/(25\rho^2)$ and $\rho \geq 5$. Let s be a sample. Let W_s be the width of candidate(s, θ). For sufficiently large n, the following hold with probability at least $1-O(n^{-\Omega(\ln^\omega n/f_{\max})})$ throughout the variation of θ within $[-\pi/10, \pi/10]$.*

(i) $W_s \leq 0.1 f(\tilde{s})$.

(ii) $\theta_s \in [-\pi/5, \pi/5]$ *and* $\theta_s = 0$ *for some* $\theta \in [-\pi/10, \pi/10]$.

(iii) Any line, which is parallel to candidate(s, θ) and inside candidate(s, θ), intersects $S_\alpha \cap$ coarse(s) for any α exactly once.

(iv) For any S_α and for any point $p \in S_\alpha \cap$ candidate(s, θ), $\theta_s - 0.2|\theta_s| - 3W_s/f(\tilde{s}) \leq \gamma_p \leq \theta_s + 0.2|\theta_s| + 3W_s/f(\tilde{s})$.

Lemma 3.4.2 *Assume that $\delta \leq 1/(25\rho^2)$ and $\rho \geq 5$. Let s be a sample. Let H be a strip that is parallel to candidate(s, θ) and lies inside candidate(s, θ). When n is sufficiently large, for any S_α and for any two points u and v on $S_\alpha \cap H$, the following hold with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.*

(i) $d(u, v) < 3\,\mathrm{width}(H)$.

(ii) The angle between the normals at u and v is at most $9\,\mathrm{width}(H)$.

(iii) The acute angle between uv and the tangent to S_α at u is at most $5\,\mathrm{width}(H)$.

3.4.1 Normal approximation

We show that our algorithm aligns *refined*(s) approximately well with the normal at $\tilde{s}$. Our algorithm varies θ so as to minimize the height of *rectangle*(s, θ). Let θ^* denote the minimizing angle. Recall that *refined*(s) = *rectangle*(s, θ^*). Let θ_s^* denote $\theta^* +$ *angle*(*strip*(s)). We apply Lemmas 3.4.1 and 3.4.2 to show that θ_s^* is very small.

Lemma 3.4.3 *Assume that $\delta \leq 1/(25\rho^2)$ and $\rho \geq 5$. Let s be a sample. Let W_s be the width of refined(s). For sufficiently large n, $|\theta_s^*| \leq 23 W_s$ with probability at least $1 - O(n^{\Omega(\ln^\omega n/f_{\max})})$.*

Proof. We rotate the plane such that $candidate(s, \theta^*)$ is vertical. Suppose that $|\theta_s^*| > 23W_s$. We first assume that Lemmas 3.3.1, 3.3.2, 3.3.3, 3.4.1, and 3.4.2 hold deterministically and show that a contradiction arises with probability at least $1 - O(n^{\Omega(\ln^\omega n/f_{\max})})$. The contradiction is that we can rotate $candidate(s, \theta^*)$ slightly to reduce its height further. Since these lemmas hold with probability at least $1 - O(n^{\Omega(\ln^\omega n/f_{\max})})$, we can then conclude that $|\theta_s^*| > 23W_s$ occurs with probability at most $O(n^{\Omega(\ln^\omega n/f_{\max})})$.

Without loss of generality, we assume that $\theta_s^* > 0$. That is, the upward normal at s points to the left. Also, we assume that $S_\delta^- \cap coarse(s)$ lies below $S_\delta^+ \cap coarse(s)$. Let L be the left boundary line of $candidate(s, \theta^*)$. By Lemma 3.4.1(iii), L intersects $S_\delta^- \cap coarse(s)$ exactly once. We use p to denote the point $L \cap S_\delta^- \cap coarse(s)$. We first prove a general claim which will be useful later.

> CLAIM 1 *Orient space such that $candidate(s, \theta)$ is vertical. If $\theta_s > 23W_s$, then for any α, $S_\alpha \cap candidate(s, \theta)$ increases strictly from left to right.*
>
> *Proof.* Take any point $z \in S_\alpha \cap candidate(s, \theta)$. By Lemma 3.4.1(iv), $\gamma_z \geq 0.8\theta_s - 3W_s$, which is positive as $\theta_s \geq 23W_s$ by assumption. Therefore, the upward normal at z points to the left, so the slope of the tangent to S_α at z is positive. □

We highlight the proof strategy before giving the details. If $\theta_s > 23W_s$, by Claim 1, both S_δ^- and S_δ^+ increase from left to right inside $candidate(s, \theta)$. Then we divide $candidate(s, \theta^*)$ into three smaller slabs of equal width in left to right order, and show that the lower side of $rectangle(s, \theta^*)$ intersects S_δ^- at a point a inside the leftmost slab. Similarly, the upper side of $rectangle(s, \theta^*)$ intersects S_δ^+ at a point b inside the rightmost slab. Since both S_δ^- and S_δ^+ increase from left to right, this allows us to rotate $rectangle(s, \theta^*)$ around a and b in the anti-clockwise direction to reduce its height. This contradicts the minimality of the height of $rectangle(s, \theta^*)$. We give the details in the following.

We first prove that the lower side of $rectangle(s, \theta^*)$ intersects S_δ^- within the leftmost slab. Let h and m be the constants in Lemma 3.3.1. Let $k = h/3240$. Let

H_1 be the slab inside $candidate(s, \theta^*)$ such that H_1 is bounded by L on the left and width$(H_1) = W_s/3$. Let H be the slab inside $candidate(s, \theta^*)$ that is bounded by L on the left and has width $30\lambda_k\sqrt{f(\tilde{s})}$. Refer to Figure 3.8. Since radius$(initial(s)) \leq$

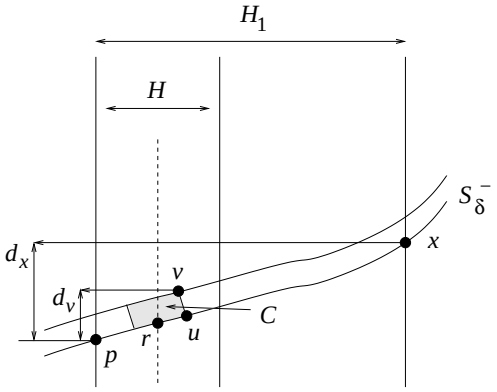

Figure 3.8: Illustration for Lemma 3.4.3.

$\psi_m\sqrt{f(\tilde{s})}$, radius$(initial(s)) < 1$ for sufficiently large n. So $\sqrt{\text{radius}(initial(s))} >$ radius$(initial(s))$. Since $W_s = \min\{\sqrt{\text{radius}(initial(s))}, \frac{\text{radius}(coarse(s))}{3}\}$, we have $W_s \geq$ radius$(initial(s))/3 \geq \lambda_h\sqrt{f(\tilde{s})}/9$. Thus

$$\text{width}(H) = 30\lambda_k\sqrt{f(\tilde{s})} = \frac{\lambda_h\sqrt{f(\tilde{s})}}{108} \leq \frac{W_s}{12}. \tag{3.3}$$

Thus, H lies inside H_1. By Lemma 3.4.1(iii), S_δ^- crosses H completely. Let r be the intersection point between S_δ^- and the center line of H. Take the $(\lambda_k/\sqrt{f_{\max}})$-grid in which $\tilde{r}$ is the first cut point. Let C be the $(\lambda_k/\sqrt{f_{\max}})$-cell such that C contains r and C lies between the normal segments at $\tilde{r}$ and the second cut point. The distance from r to the boundary of H is $15\lambda_k\sqrt{f(\tilde{s})}$. By Lemma 2.3.2, the diameter of C is at most $14\lambda_k f(\tilde{r})/\sqrt{f_{\max}} \leq 14\lambda_k\sqrt{f(\tilde{r})}$. Since $f(\tilde{r}) \leq 1.1f(\tilde{s})$ by Lemma 3.3.4(v), the diameter of C is less than $15\lambda_k\sqrt{f(\tilde{s})}$. It follows that C lies inside H.

Let u be the rightmost vertex of C on S_δ^-. Let v be the vertex of C different from u on the normal segment at u. Let x be the intersection point between S_δ^- and the right boundary line of H_1. We are to prove that x lies above C. Since C is non-empty with very high probability, the lower side of $rectangle(s, \theta^*)$ should intersect S_δ^- inside H_1 at a point below x then.

By Claim 1, v is the highest point in C and x is the highest point on $S_\delta^-(p, x)$. Let d_v and d_x be the height of v and x from p, respectively. Let ϕ be the acute

angle between pu and the horizontal line through p. Since ϕ is at most the sum of γ_p and the angle between pu and the tangent at p, by Lemma 3.4.2(iii), we have $\phi \leq \gamma_p + 5\,\mathrm{width}(H)$. By Lemma 3.4.2(i), $d(p,u) \leq 3\,\mathrm{width}(H)$. Observe that $d_v \leq d(p,u) \cdot \sin\phi + d(u,v)$. So $d_v < 3\phi\,\mathrm{width}(H) + 2\lambda_k\delta < 3\gamma_p\,\mathrm{width}(H) + 15\mathrm{width}(H)^2 + 2\lambda_k\delta$. By (3.3), we get $d_v < W_s\gamma_p/4 + 5W_s^2/48 + 2\lambda_k\delta$. We bound $2\lambda_k\delta$ as follows. Recall that $W_s = \min\{\sqrt{\mathrm{radius}(\mathit{initial}(s))}, \mathrm{radius}(\mathit{coarse}(s))/3\}$. If $W_s = \sqrt{\mathrm{radius}(\mathit{initial}(s))}$, by Lemma 3.3.1, $W_s \geq \sqrt{\lambda_h/3}f(\tilde{s})^{1/4} \geq \sqrt{\lambda_h/3}$. So $2\lambda_k\delta < 2\lambda_k = \lambda_h/1620 < 0.002W_s^2$. If $W_s = \mathrm{radius}(\mathit{coarse}(s))/3$, by Lemmas 3.3.1 and 3.3.3, $W_s \geq 2\sqrt{\rho}\delta/3$ and $W_s \geq \lambda_h\sqrt{f(\tilde{s})}/9 \geq \lambda_h/9$. We get $\lambda_k = \lambda_h/3240 \leq W_s/360$ and $2\delta \leq 3W_s/\sqrt{\rho} \leq 3W_s/\sqrt{5}$, so $2\lambda_k\delta < 0.004W_s^2$. We conclude that

$$d_v < \frac{W_s\gamma_p}{4} + 0.2W_s^2.$$

We observe that px is parallel to the tangent at some point z on $S_\delta^-(p,x)$. Then by Lemma 3.4.2 (ii), $\gamma_z \geq \gamma_p - 9\,\mathrm{width}(H_1) = \gamma_p - 3W_s$. Since $d_x = \mathrm{width}(H_1) \cdot \tan\gamma_z = (W_s/3) \cdot \tan\gamma_z$, we get

$$d_x \geq \frac{W_s\gamma_z}{3} \geq \frac{W_s\gamma_p}{3} - W_s^2.$$

Since $\theta_s^* > 23W_s$ by our assumption, Lemma 3.4.1(iv) implies that $\gamma_p \geq 0.8\theta_s^* - 3W_s > 15W_s$. Therefore, $d_x - d_v > W_s\gamma_p/12 - 1.2W_s^2 > 0$. It follows that x lies above C.

Since C is a $(\lambda_k/\sqrt{f_{\max}})$-cell, by Lemma 2.3.6(i), C contains some sample with probability at least $1 - n^{\Omega(\ln^\omega n/f_{\max})}$. Thus, the lower side of $\mathit{rectangle}(s,\theta^*)$ lies below x with probability at least $1 - n^{\Omega(\ln^\omega n/f_{\max})}$. On the other hand, the lower side of $\mathit{rectangle}(s,\theta^*)$ cannot lie below $S_\delta^- \cap H_1$, otherwise it could be raised to reduce the height of $\mathit{rectangle}(s,\theta^*)$, a contradiction. So the lower side of $\mathit{rectangle}(s,\theta^*)$ intersects $S_\delta^- \cap H_1$ at some point a. See the left figure in Figure 3.9.

Let H_2 be the slab inside $\mathit{candidate}(s,\theta^*)$ such that H_2 is bounded by the right boundary line of $\mathit{candidate}(s,\theta^*)$ on the right and $\mathrm{width}(H_2) = W_s/3$. By a symmetric argument, we can prove that the upper side of $\mathit{rectangle}(s,\theta^*)$ intersects $S_\delta^+ \cap H_2$ at a point b.

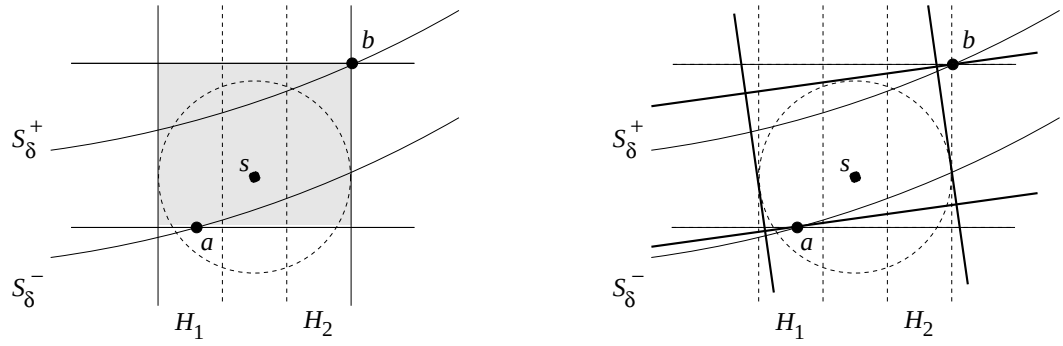

Figure 3.9: In the right figure, the middle bold rectangle is the obtained by a slight anti-clockwise rotation. Its height is smaller than that of the middle dashed rectangle.

Consider an angle θ that is slightly less than θ^*. As shown in the right figure in Figure 3.9, this is equivalent to rotating the candidate neighborhood in the anti-clockwise direction. By Lemma 3.4.1(ii), θ_s can reach zero during the variation of θ. Thus, as $\theta_s^* > 0$, decreasing θ from θ^* is legal. Moreover, as $\theta_s^* > 23W_s$, the small rotation keeps θ_s greater than $23W_s$. Correspondingly, we rotate the lower and upper sides of $\mathit{rectangle}(s, \theta^*)$ around a and b, respectively, to obtain a rectangle R. Orient the plane such that the new candidate neighborhood becomes vertical. By Claim 1, S_δ^- increases strictly from left to right, so S_δ^- crosses the lower side of R at most once at a from below to above. Similarly, S_δ^+ crosses the upper side of R at most once at b from below to above. This implies that R contains all the samples inside the new candidate neighborhood . Since a is on the left of b and below b, the anti-clockwise rotation makes the height of R strictly less than the height of $\mathit{rectangle}(s, \theta^*)$. This contradicts the assumption that the height of $\mathit{rectangle}(s, \theta^*)$ is already the minimum possible. □

3.4.2 Pointwise convergence

Once $\mathit{refined}(s)$ is aligned well with the normal at $\tilde{s}$, it is intuitively true that the center point of $\mathit{refined}(s)$ should lie very close to $\tilde{s}$. The following lemma proves this formally.

Lemma 3.4.4 *Assume that* $\delta \leq 1/(25\rho^2)$ *and* $\rho \geq 5$. *Let* s *be a sample. Let* W_s *be the width of* $refined(s)$. *For sufficiently large* n, *the distance between the center point* s^* *of* $refined(s)$ *and* $\tilde{s}$ *is at most* $(138\delta + 3)W_s$ *with probability at least* $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.

Proof. We first assume that Lemmas 3.3.1, 3.3.2, 3.3.3, 3.4.1, 3.4.2, and 3.4.3 hold deterministically and show that the lemma is true with probability at least $1 - O(n^{\Omega(\ln^\omega n/f_{\max})})$. As these lemmas hold with probability at least $1-O(n^{\Omega(\ln^\omega n/f_{\max})})$, the lemma follows.

Assume that s lies on S_α^+, the normal at $\tilde{s}$ is vertical, and $S_\delta^+ \cap coarse(s)$ is above $S_\delta^- \cap coarse(s)$. Let r_d (resp., r_u) be the ray that shoots downward (resp., upward) from s and makes an angle θ_s^* with the vertical. Let x and y be the points on S_δ^+ and S hit by r_u and r_d respectively. Let z be the point on S_δ^- hit by r_d. Let s_1 be the point on S_δ^- such that $\tilde{s_1} = \tilde{s}$. Without loss of generality, we assume that $\theta_s^* \geq 0$. Refer to Figure 3.10.

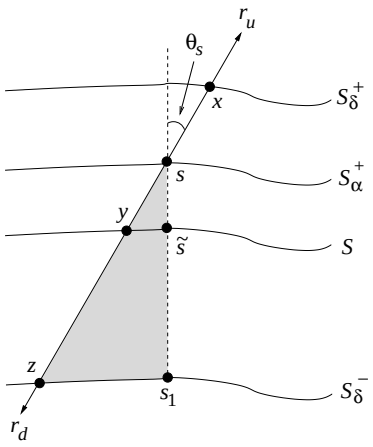

Figure 3.10: Illustration for Lemma 3.4.4.

Our strategy for bounding $d(\tilde{s}, s^*)$ is as follows. By triangle inequality, $d(\tilde{s}, s^*) \leq d(s^*, y) + d(\tilde{s}, y)$. Thus it suffices to bound $d(s^*, y)$ and $d(\tilde{s}, y)$. While $d(\tilde{s}, y)$ can be bounded directly, a few intermediate steps are needed to bound $d(s^*, y)$. If the upper and lower sides of $refined(s)$ pass through x and z, respectively, then $d(s^*, y)$ is just the distance between the midpoint of xz and y. Then we consider the cases that the upper and lower sides of $refined(s)$ do not pass through x and z, and bound

the maximum displacement of s^* from the midpoint of xz. This yields the bound on $d(s^*, y)$. We give the details in the following.

We first bound the distance between y and the midpoint of xz. By Lemma 3.3.4(iv), the acute angle between s_1z and the tangent at s_1 (the horizontal) is at most $\sin^{-1}(0.03)$. It follows that $\angle ss_1z \leq \pi/2 + \sin^{-1}(0.03)$. So $\angle szs_1 = \pi - \theta_s^* - \angle ss_1z \geq \pi/2 - \theta_s^* - \sin^{-1}(0.03)$, which is greater than 0.9 as $\theta_s^* \leq \pi/5$ by Lemma 3.4.1(ii). By applying sine law to the shaded triangle in Figure 3.10, we get

$$d(s_1, z) = \frac{d(s, s_1) \cdot \sin\theta_s^*}{\sin\angle szs_1} \leq \frac{(\delta+\alpha)\theta_s^*}{\sin(0.9)} < 2(\delta+\alpha)\theta_s^*. \tag{3.4}$$

Similarly, we get

$$d(\tilde{s}, y) = \frac{d(s, \tilde{s}) \cdot \sin\theta_s^*}{\sin\angle sy\tilde{s}} \leq \frac{\alpha\theta_s^*}{\sin(0.9)} < 2\alpha\theta_s^*. \tag{3.5}$$

By triangle inequality, $d(s, s_1) - d(s_1, z) \leq d(s, z) \leq d(s, s_1) + d(s_1, z)$. Then (3.4) yields

$$(\delta+\alpha) - 2(\delta+\alpha)\theta_s^* \leq d(s, z) \leq (\delta+\alpha) + 2(\delta+\alpha)\theta_s^*. \tag{3.6}$$

We can use a similar argument to show that

$$(\delta-\alpha) - 2(\delta-\alpha)\theta_s^* \leq d(s, x) \leq (\delta-\alpha) + 2(\delta-\alpha)\theta_s^*, \tag{3.7}$$

$$\alpha - 2\alpha\theta_s^* \leq d(s, y) \leq \alpha + 2\alpha\theta_s^*. \tag{3.8}$$

Let d_x and d_y be the distances from the midpoint of xz to x and y, respectively. Since $d(x, z) = d(s, x) + d(s, z)$, by (3.6) and (3.7), we get $2\delta - 4\delta\theta_s^* \leq d(x, z) \leq 2\delta + 4\delta\theta_s^*$. Therefore, $\delta - 2\delta\theta_s^* \leq d_x \leq \delta + 2\delta\theta_s^*$. Since $d(x, y) = d(s, x) + d(s, y)$, by (3.7) and (3.8), we get $\delta - 2\delta\theta_s^* \leq d(x, y) \leq \delta + 2\delta\theta_s^*$. We conclude that

$$d_y = |d_x - d(x, y)| \leq 4\delta\theta_s^*. \tag{3.9}$$

Second, we bound the displacement of s^* from the midpoint of xz. There are two cases.

Case 1: the upper side of *refined*(s) lies above x. The upper side of *refined*(s) must intersect $S_\delta^+ \cap$ *candidate*(s, θ^*) at some point v, otherwise we could lower it to reduce the height of *refined*(s), a contradiction. Since $d(x, v) \leq 3W_s$ by Lemma 3.4.2(i), the distance between x and the upper side of *refined*(s) is at most $3W_s$.

Case 2: the upper side of *refined*(s) lies below x. Let h be the constant in Lemma 3.3.1. Let $k = h/270$. Take the $(\lambda_k/\sqrt{f_{\max}})$-grid in which $\tilde{x}$ is the first cut point. Let C be the cell such that C contains x and C lies between the normal segments at $\tilde{x}$ and the second cut point.

We claim that C lies inside *candidate*(s, θ^*). Since radius(*initial*(s)) $\leq \psi_m\sqrt{f(\tilde{s})}$, we have radius(*initial*(s)) < 1 for sufficiently large n. Thus we get $\sqrt{\text{radius}(\textit{initial}(s))} > \text{radius}(\textit{initial}(s))$. So

$$\begin{aligned} W_s &= \min\{\sqrt{\text{radius}(\textit{initial}(s))}, \text{radius}(\textit{coarse}(s))/3\} \\ &\geq \text{radius}(\textit{initial}(s))/3, \end{aligned}$$

which is at least $\lambda_h\sqrt{f(\tilde{s})}/9$. By Lemma 2.3.2, the diameter of C is at most $14\lambda_k f(\tilde{x})/\sqrt{f_{\max}} \leq 14\lambda_k\sqrt{f(\tilde{x})}$. Since $f(\tilde{x}) \leq 1.1f(\tilde{s})$ by Lemma 3.3.4(v), the diameter of C is less than $15\lambda_k\sqrt{f(\tilde{s})}$. Since $W_s \geq \lambda_h\sqrt{f(\tilde{s})}/9 = 30\lambda_k\sqrt{f(\tilde{s})}$, C must lie inside *candidate*(s, θ^*).

Since C is a $(\lambda_k/\sqrt{f_{\max}})$-cell, by Lemma 2.3.6(i), C contains some sample with probability at least $1 - n^{-\Omega(\ln^\omega n/f_{\max})}$. Thus, the upper side of *refined*(s) cannot lie below C. It follows that the distance between x and the upper side of *refined*(s) is at most the diameter of C, which has been shown to be less than $W_s/2$.

Hence, the position of the upper side of *refined*(s) may cause s^* to be displaced from the midpoint of xz by a distance of at most $3W_s/2$. The position of the lower side of *refined*(s) has the same effect. So the distance between s^* and the midpoint of xz is at most $3W_s$. It follows that $d(s^*, y) \leq d_y + 3W_s$. By (3.9), we get $d(s^*, y) \leq 4\delta\theta_s^* + 3W_s$. Starting with triangle inequality, we obtain

$$\begin{aligned} d(\tilde{s}, s^*) &\leq d(s^*, y) + d(\tilde{s}, y) \\ &\leq 4\delta\theta_s^* + 3W_s + d(\tilde{s}, y) \\ &\overset{(3.5)}{\leq} 6\delta\theta_s^* + 3W_s. \end{aligned}$$

Since $\theta_s^* \leq 23W_s$ by Lemma 3.4.3, we conclude that $d(\tilde{s}, s^*) \leq (138\delta + 3)W_s$. □

3.5 Homeomorphism

In this section, we prove more convergence properties which lead to the proof that the output curve of the NN-crust algorithm is homeomorphic to S. For each sample s, we use s^* to denote the center point of *refined*(s). We briefly review the processing of the center points. We first sort the center points in decreasing order of the widths of their corresponding refined neighborhoods. Then we scan the sorted list to select a subset of center points. When the current center point s^* is selected, we delete all center points p^* from the sorted list such that $d(p^*, s^*) \leq \text{width}(\textit{refined}(s))^{1/3}$.

In the end, we call two selected center points s^* and t^* *adjacent* if $S(\tilde{s}, \tilde{t})$ or $S(\tilde{t}, \tilde{s})$ does not contain $\tilde{u}$ for any other selected center point u^*. We use G to denote the polygonal curve that connects adjacent selected center points. Note that the degree of every vertex in G is exactly two. Clearly, if we connect $\tilde{s}$ and $\tilde{t}$ for every pair of adjacent selected center points s^* and t^*, we obtain a polygonal curve G' that is homeomorphic to S. Our goal is to show that the output curve of the NN-crust algorithm is exactly G. Since there is a bijection between G and G', the homeomorphism result follows.

Throughout this section, we assume that $\text{width}(\textit{initial}(s)) < 1$ for any sample s, which is true for sufficiently large n. There are a few consequences. First, it implies that $\sqrt{\text{radius}(\textit{initial}(s))} \geq \text{radius}(\textit{initial}(s))$. Second, since

$$\text{width}(\textit{refined}(s)) = \min\{\sqrt{\text{radius}(\textit{initial}(s))}, \text{radius}(\textit{coarse}(s))/3\},$$

we get $\text{width}(\textit{refined}(s)) \leq \sqrt{\text{radius}(\textit{initial}(s))} < 1$. This implies for any constants $a > b > 0$, $\text{width}(\textit{refined}(s))^a < \text{width}(\textit{refined}(s))^b$. Lastly, $\text{width}(\textit{refined}(s)) \geq \text{radius}(\textit{initial}(s))/3$.

We need the technical results Lemmas 3.5.1–3.5.6. The proofs of Lemmas 3.5.1, 3.5.3, 3.5.4, and 3.5.5 are given in the appendix.

Lemma 3.5.1 *There exists a constant $\mu_1 > 0$ such that when n is sufficiently large, for any two center points p^* and q^*, if $d(\tilde{p}, \tilde{q}) \leq f(\tilde{p})/2$, then $W_q \leq \mu_1 f(\tilde{p})\sqrt{W_p}$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.*

Lemma 3.5.2 *Let p^* and q^* be two selected center points. Then*

$$d(p^*, q^*) > \max\{W_p^{1/3}, W_q^{1/3}\}.$$

Proof. Assume without loss of generality that p^* was selected before q^*. Since q^* was selected subsequently, q^* was not eliminated by the selection of p^*. Thus, $d(p^*, q^*) > W_p^{1/3} \geq W_q^{1/3}$. □

Lemma 3.5.3 *When n is sufficiently large, for any two center points x^* and y^* such that $d(\tilde{x}, \tilde{y}) \leq f(\tilde{y})/2$ and $d(x^*, y^*) \geq W_y^{1/3}$, the acute angle between x^*y^* and $\tilde{x}\tilde{y}$ is $O(f(\tilde{y})W_y^{1/6})$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.*

Lemma 3.5.4 *When n is sufficiently large, for any three center points x^*, y^*, and z^* such that $\tilde{y} \in S(\tilde{x}, \tilde{z})$, $d(\tilde{x}, \tilde{z}) \leq \max\{f(\tilde{x})/5, f(\tilde{z})/5\}$, $d(x^*, y^*) \geq W_y^{1/3}$, and $d(y^*, z^*) \geq W_y^{1/3}$, the angle $\angle x^*y^*z^*$ is obtuse with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.*

Lemma 3.5.5 *There exists a constant $\mu_2 > 0$ such that when n is sufficiently large, for any edge e in G connecting two center points p^* and q^*, $\text{length}(e) \leq \mu_2 f(\tilde{p})W_p^{1/3} + \mu_2 f(\tilde{q})W_q^{1/3}$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.*

Lemma 3.5.6 *When n is sufficiently large, for any two selected center points p^* and q^* such that p^* and q^* are not adjacent in G and $d(p^*, q^*) \leq f(\tilde{p})/5$, there is an edge e in G incident to p^* such that the angle between e and p^*q^* is acute and $\text{length}(e) < d(p^*, q^*)$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$.*

Proof. Since p^* and q^* are not adjacent in G, there is some selected center point u^* adjacent to p^* such that $\tilde{u}$ lies on $S(\tilde{p}, \tilde{q})$ or $S(\tilde{q}, \tilde{p})$, say $S(\tilde{p}, \tilde{q})$. By Lemma 3.5.2, $d(p^*, u^*) > W_u^{1/3}$ and $d(q^*, u^*) > W_u^{1/3}$. By Lemma 3.5.4, the angle $\angle p^*u^*q^*$ is obtuse with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$. It follows that $\angle u^*p^*q^*$ is acute and $d(p^*, u^*) < d(p^*, q^*)$. □

We apply the above technical lemmas to show that the output curve of the NN-crust algorithm is exactly G. Then this allows us to show that the output curve is homeomorphic to the underlying smooth closed curve.

Lemma 3.5.7 *For sufficiently large n, the output curve obtained by running the NN-crust algorithm on the selected center points is exactly G with probability at least $1 - O(n^{-\Omega(\frac{\ln^{\omega} n}{f_{\max}} - 1)})$.*

Proof. We first prove the lemma assuming that Lemmas 3.4.4, 3.5.4, 3.5.5, and 3.5.6 hold deterministically. We will discuss the probability bound later.

Let p^* be a selected center point. Let p^*u^* and p^*v^* be the edges of G incident to p^*. Without loss of generality, we assume that $\tilde{p}$ lies on $S(\tilde{u}, \tilde{v})$. By Lemma 3.5.2, $d(p^*, u^*) > W_p^{1/3}$ and $d(p^*, v^*) > W_p^{1/3}$.

Let $k = 138\delta + 3$. By Lemmas 3.4.4 and 3.5.5, $d(\tilde{p}, \tilde{u}) \leq d(\tilde{p}, p^*) + d(\tilde{u}, u^*) + d(p^*, u^*) \leq kW_p + kW_u + \mu_2 f(\tilde{p}) W_p^{1/3} + \mu_2 f(\tilde{u}) W_u^{1/3}$, which is less than $(f(\tilde{p}) + f(\tilde{u}))/30$ for sufficiently large n. The Lipschitz condition implies that

$$0.9 f(\tilde{p}) < f(\tilde{u}) < 1.1 f(\tilde{p}).$$

So we get

$$d(\tilde{p}, \tilde{u}) \leq \frac{f(\tilde{p}) + f(\tilde{u})}{30} < 0.07 f(\tilde{p}), \qquad d(p^*, u^*) \leq \frac{f(\tilde{p}) + f(\tilde{u})}{30} < 0.07 f(\tilde{p}).$$

Similarly, we can show that

$$d(\tilde{p}, \tilde{v}) < 0.07 f(\tilde{p}), \qquad d(p^*, v^*) < 0.07 f(\tilde{p}).$$

Let p^*q^* be an edge computed by the NN-crust algorithm when it processes the vertex p^*. Assume to the contrary that p^*q^* is not an edge in G. If p^*q^* is computed in step 1 of the NN-crust algorithm, then q^* is the nearest neighbor of p^*. So $d(p^*, q^*) \leq d(p^*, u^*) < 0.07 f(\tilde{p})$. By Lemma 3.5.6, there is another edge e in G such that $\text{length}(e) < d(p^*, q^*)$, a contradiction. Suppose that p^*q^* is computed in step 2 of the NN-crust algorithm. As we have just shown, the step 1 of the NN-crust algorithm already outputs an edge, say p^*u^*, of G where u^* is the nearest

neighbor of p^*. Observe that $d(\tilde{u}, \tilde{v}) \leq d(\tilde{p}, \tilde{u}) + d(\tilde{p}, \tilde{v}) < 0.14f(\tilde{p}) < 0.2f(\tilde{u})$. By Lemma 3.5.4, $\angle u^*p^*v^*$ is obtuse. By the NN-crust algorithm, $\angle u^*p^*q^*$ is also obtuse. Since the NN-crust algorithm prefers p^*q^* to p^*v^*, $d(p^*, q^*) \leq d(p^*, v^*) < 0.07f(\tilde{p})$. By Lemma 3.5.6, G has an edge e incident to p^* that is shorter than p^*q^* (p^*v^* too) and makes an acute angle with p^*q^*. The edge e is not p^*v^* as e is shorter than p^*v^*. The edge e is not p^*u^* as $\angle u^*p^*q^*$ is obtuse. But then the degree of p in G is at least three, a contradiction.

We have shown that each output edge belongs to G. Since the NN-crust algorithm guarantees that each vertex in the output curve has degree at least two, the output curve and G have the same number of edges. So the output curve is exactly G.

Since Lemmas 3.4.4, 3.5.4, 3.5.5, and 3.5.6 hold with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$, the output edges incident to p^* are edges of G with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$. Since there are $O(n)$ output vertices, the probability that this holds for all vertices is at least $1 - O(n^{-\Omega(\frac{\ln^\omega n}{f_{\max}} - 1)})$. □

Corollary 3.5.1 *For sufficiently large n, the output curve obtained by running the NN-crust algorithm on the selected center points is homeomorphic to the underlying smooth closed curve with probability at least $1 - O(n^{-\Omega(\frac{\ln^\omega n}{f_{\max}} - 1)})$.*

Proof. We have shown that the output curve is G. Let G' be the curve obtained by connecting $\tilde{p}$ and $\tilde{q}$ for each edge p^*q^* of G. G' is homeomorphic to the underlying smooth closed curve as p^* and q^* are adjacent in G. Clearly, G is homeomorphic to G' as there is a bijection between the edges of G and G'. □

3.6 Finale

We make use of the lemmas in the previous subsections to prove the key result of this paper, stated as the Main Theorem in Section 3.2.

Proof of the Main Theorem. First of all, for any sample s, let W_s denote the width of *refined*(s). By construction, $W_s \leq \sqrt{\text{radius}(\textit{initial}(s))}$. By Lemma 3.3.1, $\text{radius}(\textit{initial}(s)) = O((\frac{\ln^{1+\omega} n}{n})^{1/4} f(\tilde{s})^{1/2})$. Thus $W_s = O((\frac{\ln^{1+\omega} n}{n})^{1/8} f(\tilde{s})^{1/4})$.

By Lemma 3.4.4, as n tends to ∞, for each output vertex s^*, $d(s^*, \tilde{s}) = O(W_s)$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n / f_{\max})})$. Since there are $O(n)$ output vertices, the distance bounds hold simultaneously with probability at least $1 - O(n^{-\Omega(\frac{\ln^\omega n}{f_{\max}} - 1)})$. Next, we analyze the angular differences between the tangents of the smooth closed curve and the output curve.

Let r^*s^* be an output edge. By Lemma 3.5.5, with probability at least $1 - O(n^{-\Omega(\ln^\omega n / f_{\max})})$, we have

$$d(r^*, s^*) \leq \mu_2 f(\tilde{r}) W_r^{1/3} + \mu_2 f(\tilde{s}) W_s^{1/3}. \tag{3.10}$$

Let $k = 138\delta + 3$. Using the above, the triangle inequality, and Lemma 3.4.4, we get

$$\begin{aligned} d(\tilde{r}, \tilde{s}) &\leq d(\tilde{r}, r^*) + d(\tilde{s}, s^*) + d(r^*, s^*) && (3.11) \\ &\leq k W_r + k W_s + \mu_2 f(\tilde{r}) W_r^{1/3} + \mu_2 f(\tilde{s}) W_s^{1/3}. && (3.12) \end{aligned}$$

By (3.10), $d(r^*, s^*) < f(\tilde{r})/5 + f(\tilde{s})/5$ for sufficiently large n. The Lipschitz condition implies that $f(\tilde{r}) < 1.5 f(\tilde{s})$. So $d(r^*, s^*) < f(\tilde{s})/2$. Thus, Lemma 3.5.1 applies and yields $W_r \leq \mu_1 f(\tilde{s}) \sqrt{W_s}$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n / f_{\max})})$. Substituting into (3.12), we conclude that

$$d(\tilde{r}, \tilde{s}) \leq \mu_3 f(\tilde{s})^{4/3} W_s^{1/6}, \tag{3.13}$$

for some constant $\mu_3 > 0$.

Let θ be the angle between $\tilde{r}\tilde{s}$ and the tangent at $\tilde{s}$. By Lemma 2.2.2(ii), we have $\theta \leq \sin^{-1} \frac{\mu_3 f(\tilde{s})^{1/3} W_s^{1/6}}{2}$. Let θ' be the acute angle between r^*s^* and $\tilde{r}\tilde{s}$. By (3.13), $d(\tilde{r}, \tilde{s}) \leq f(\tilde{s})/2$ for sufficiently large n. Thus, by Lemma 3.5.3, $\theta' = O(f(\tilde{s}) W_s^{1/6})$ with probability at least $1 - O(n^{-\Omega(\ln^\omega n / f_{\max})})$ for sufficiently large n. We conclude that the angle between r^*s^* and the tangent at $\tilde{s}$, which is upper bounded by $\theta + \theta'$, is $O(f(\tilde{s}) W_s^{1/6})$. Since there are $O(n)$ output edges, the angular difference bounds hold simultaneously with probability at least $1 - O(n^{-\Omega(\frac{\ln^\omega n}{f_{\max}} - 1)})$.

The output curve is homeomorphic to the smooth closed curve by Corollary 3.5.1. □

3.7 Summary

We have presented an algorithm to reconstruct polygonal closed curves from noisy samples drawn from a set of smooth closed curves. The output polygonal reconstruction converges to the original curve with probability approaching to 1 as n increases. Although we have assumed that there is only one smooth closed curve in our analysis for notational simplicity, the analysis can be carried over to the general case. A straightforward implementation of our algorithm takes $O(n^3)$ time. We view the analysis as our major contribution as it is the first result that deals with faithful curve reconstruction from noisy samples. Since the analysis is already quite involved, we did not spend much effort in looking for a faster algorithm.

3.8 Appendix

Proof of Lemma 3.3.4

A straightforward calculation shows (i).

If $S_\alpha \cap \mathit{coarse}(s)$ consists of more than one connected component, the medial axis of S_α intersects the interior of $\mathit{coarse}(s)$. Since S and S_α have the same medial axis, the distance from $\tilde{s}$ to the medial axis is at most $2\,\mathrm{radius}(\mathit{coarse}(s)) \leq 2(5\rho\delta + \psi_m\sqrt{f(\tilde{s})}) \leq 2(5\rho\delta + \psi_m)f(\tilde{s}) < f(\tilde{s})$ by (i), a contradiction. This proves (ii).

Let s_1 be the point on S_α such that $\tilde{s_1} = \tilde{s}$. The distance $d(s_1, x) \leq d(s, x) + d(s, s_1) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})} + 2\delta \leq (5\rho\delta + \psi_m + 2\delta)f(\tilde{s})$. By Lemma 2.2.3, the angle between the normals at s_1 and x is at most $2\sin^{-1}\frac{d(s_1,x)}{(1-\delta)f(\tilde{s})} \leq 2\sin^{-1}\frac{5\rho\delta+\psi_m+2\delta}{(1-\delta)} \leq 2\sin^{-1}(0.06)$ by (i). This proves (iii).

By Lemma 2.2.2(ii), $x \in \mathit{cocone}(s_1, 2\sin^{-1}\frac{d(s_1,x)}{2(1-\delta)f(\tilde{s})}) \subseteq \mathit{cocone}(s_1, 2\sin^{-1}(0.03))$. This proves (iv).

The distance $d(\tilde{s}, \tilde{x}) \leq d(s, \tilde{s}) + d(s, x) + d(x, \tilde{x}) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})} + 2\delta \leq (5\rho\delta + \psi_m + 2\delta)f(\tilde{s}) < 0.1f(\tilde{s})$. Then the Lipschitz condition implies (v).

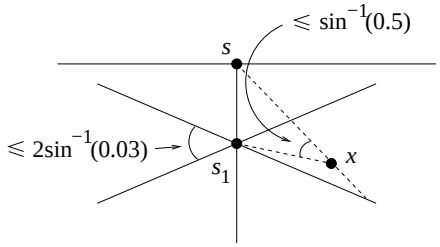

Figure 3.11: Illustration for Lemma 3.3.4.

Consider (vi). Refer to Figure 3.11. Assume that the tangent at s is horizontal. By sine law, $\sin\angle sxs_1 = \frac{d(s,s_1)\cdot\sin\angle ss_1x}{d(s,x)} \leq \frac{2\delta}{\mathrm{radius}(\mathit{coarse}(s))}$ as $d(s, s_1) \leq 2\delta$ and $d(s, x) = \mathrm{radius}(\mathit{coarse}(s))$. Since $\mathrm{radius}(\mathit{coarse}(s)) \geq 2\sqrt{\rho}\delta$ and $\rho \geq 5$, we have $\angle sxs_1 \leq \sin^{-1}\frac{1}{\sqrt{\rho}} < \sin^{-1}(0.5)$. By (iv), $\angle s_1sx \geq \pi - \angle sxs_1 - (\pi/2 + \sin^{-1}(0.03)) > \pi/2 - \sin^{-1}(0.5) - \sin^{-1}(0.03)$. Thus, the horizontal distance between s and x is equal to $d(s, x) \cdot \sin\angle s_1sx \geq d(s, x) \cdot \cos(\sin^{-1}(0.5) + \sin^{-1}(0.03)) > 0.8 \cdot d(s, x)$.

Consider (vii). Since $y \in S_\alpha \cap \mathit{coarse}(s)$, $d(x,y) \leq 2\,\mathrm{radius}(\mathit{coarse}(s)) \leq 2(5\rho\delta + \psi_m\sqrt{f(\tilde{s})})$ which is at most $0.1f(\tilde{s})$ by (i). So Lemma 2.2.2(ii) applies and the acute angle between xy and the tangent at x is at most $\sin^{-1}\frac{d(x,y)}{2(1-\delta)f(\tilde{x})} \leq \sin^{-1}\frac{(5\rho\delta+\psi_m)f(\tilde{s})}{(1-\delta)f(\tilde{x})}$. Since $f(\tilde{x}) \geq 0.9f(\tilde{s})$ by (v) and $\delta \leq 1/(25\rho^2)$, the acute angle is less than $\sin^{-1}(1.2(5\rho\delta + \psi_m))$, which is less than $\sin^{-1}(0.06)$ by (i). □

Proof of Lemma 3.4.1

We first assume that $\max\{2\sqrt{\rho}\delta, \psi_h\sqrt{f(\tilde{s})}\} \leq \mathrm{radius}(\mathit{coarse}(s)) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$ and $\mathrm{radius}(\mathit{initial}(s)) \leq \psi_m\sqrt{f(\tilde{s})}$. We will take the probabilities of their occurrences later into consideration.

Since $W_s \leq \sqrt{\mathrm{radius}(\mathit{initial}(s))} \leq \sqrt{\psi_m} f(\tilde{s})^{1/4}$ and $\psi_m \leq 0.01$ for sufficiently large n, $W_s \leq 0.1f(\tilde{s})$. This proves (i).

By Lemma 3.3.5, for sufficiently large n, $|\mathit{angle}(\mathit{strip}(s))| \leq 4\sin^{-1}(0.06) < \pi/10$. Since $\theta \in [-\pi/10, \pi/10]$, $\theta_s = \theta + \mathit{angle}(\mathit{strip}(s)) \in [-\pi/5, \pi/5]$ and $\theta_s = 0$ for some θ. This proves (ii).

Consider (iii). Let ℓ be a line that is parallel to $\mathit{candidate}(s,\theta)$ and inside $\mathit{candidate}(s,\theta)$. We first prove that ℓ intersects S_α. Refer to Figure 3.12. Without loss of generality, assume that the normal at $\tilde{s}$ is vertical, the slope of $\mathit{candidate}(s,\theta)$ is positive, and ℓ is below s. Let s_1 and s_2 be the points on S_δ^+ and S_δ^-, respectively, such that $\tilde{s_1} = \tilde{s_2} = \tilde{s}$. Shoot two rays upward from s_1 with slopes $\pm\sin^{-1}(0.03)$. Also, shoot two rays downward from s_2 with slopes $\pm\sin^{-1}(0.03)$. Let $\mathcal{R}$ be the region inside $\mathit{coarse}(s)$ bounded by these four rays. By Lemma 3.3.4(iv), $S_\alpha \cap \mathit{coarse}(s)$ lies inside $\mathcal{R}$. Let x be the upper right vertex of $\mathcal{R}$. Let y be the right endpoint of a horizontal chord through s_1. Let L be the line that passes through x and is parallel to ℓ. Let L' be the line that passes through s and is parallel to ℓ. Let z be the point on L such that s_1z is perpendicular to L.

We claim that L' is above L and L and L' intersect both the upper and lower boundaries of $\mathcal{R}$. By Lemma 3.3.4(iv), $\angle xs_1y \leq \sin^{-1}(0.03)$, so $\angle xsy \leq 2\sin^{-1}(0.03)$. Observe that $\cos\angle s_1sy = \frac{d(s,s_1)}{d(s,y)} \leq \frac{2\delta}{\mathrm{radius}(\mathit{coarse}(s))}$. Since $\mathrm{radius}(\mathit{coarse}(s)) \geq 2\sqrt{\rho}\delta$, $\cos\angle s_1sy \leq 1/\sqrt{\rho} \leq 1/\sqrt{5}$ which implies that $\angle s_1sy > \pi/3$. Since $\angle s_1sx =$

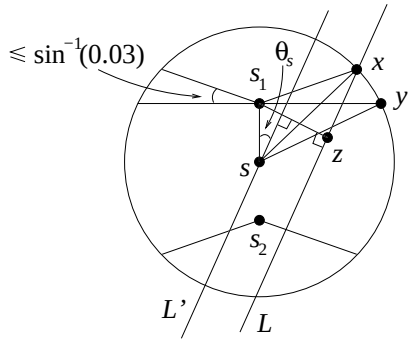

Figure 3.12: Illustration for Lemma 3.4.1(iii).

$\angle s_1sy - \angle xsy$, we get

$$\angle s_1sx \geq \pi/3 - 2\sin^{-1}(0.03) > \pi/5 \geq |\theta_s|. \tag{3.14}$$

So L' cuts through the angle between ss_1 and sx. It follows that L' is above L. Observe that L' intersects s_1x. By symmetry, L' intersects the left downward ray from s_2 too. We conclude that L and L' intersect both the upper and lower boundaries of $\mathcal{R}$.

Since $|\theta_s| \leq \pi/5$ and $\angle sxz = \angle s_1sx - |\theta_s|$, by (3.14), $\angle sxz \geq \pi/3 - 2\sin^{-1}(0.03) - \pi/5 > 0.3$. The distance from s to L is equal to $d(s,x) \cdot \sin\angle sxz > d(s,x) \cdot \sin(0.3) > 0.2 \cdot \text{radius}(\mathit{coarse}(s))$. Recall that ℓ lies below s by our assumption. The distance between ℓ and s is at most $W_s/2$ and our algorithm enforces that $W_s/2 \leq \text{radius}(\mathit{coarse}(s))/6$. So ℓ lies between L' and L. Since L and L' intersect both the upper and lower boundaries of $\mathcal{R}$, so does ℓ. It follows that ℓ must intersect $S_\alpha \cap \mathit{coarse}(s)$.

Next, we show that ℓ intersects $S_\alpha \cap \mathit{coarse}(s)$ exactly once. If not, ℓ is parallel to the tangent at some point on $S_\alpha \cap \mathit{coarse}(s)$. By Lemma 3.3.4(iii), the angle between ℓ and the vertical is at least $\pi/2 - 2\sin^{-1}(0.06) > \pi/5$, contradicting the fact that $|\theta_s| \leq \pi/5$.

Consider (iv). Let ℓ be a line that is parallel to $\mathit{candidate}(s,\theta)$ and passes through s. By (iii), ℓ intersects S_α at some point b. We first prove that $\theta_s - 0.2|\theta_s| \leq \gamma_b \leq \theta_s + 0.2|\theta_s|$. Let s_1 be the point on S_α such that $\tilde{s} = \tilde{s_1}$. Assume that the tangent at s is horizontal, s is above s_1, and b is to the left of s. Let C be the circle tangent to S_α at s_1 that lies below s_1, is centered at x, and has radius $f(\tilde{s}) - \delta$. By Lemma 2.2.1,

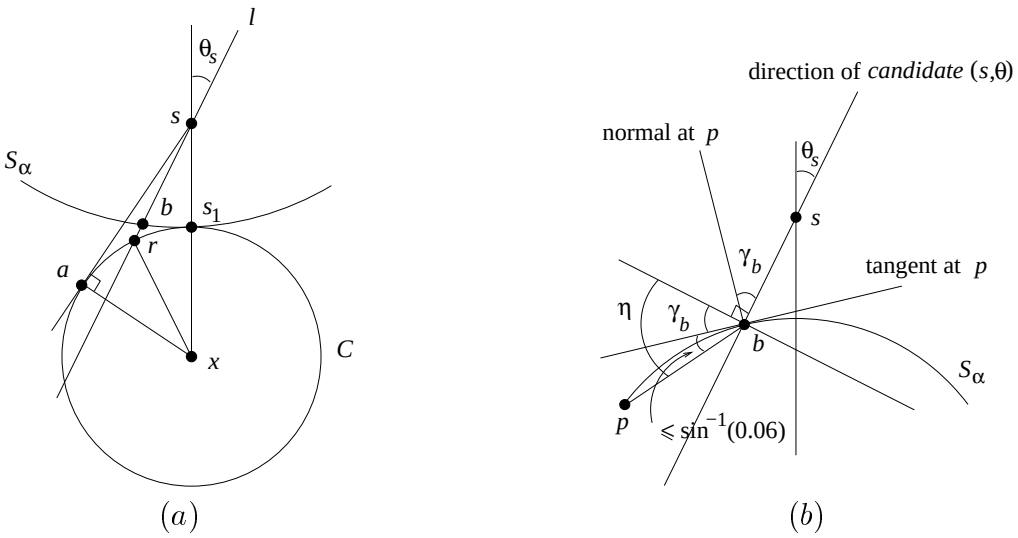

Figure 3.13: Illustration for Lemma 3.4.1(iv).

S_α does not intersect the interior of C. Refer to Figure 3.13(a). Let sa be a tangent to C that lies on the left of x. We claim that $\angle asx > |\theta_s|$. Otherwise, $d(s,x) \geq d(a,x)/\sin(\pi/5) = (f(\tilde{s}) - \delta)/\sin(\pi/5) > f(\tilde{s}) + \delta \geq d(s,x)$, a contradiction. So sb lies between sa and sx. Let sr be the extension of sb such that r lies on C. We have $d(a,s) = \sqrt{d(s,x)^2 - d(a,x)^2} \leq \sqrt{(f(\tilde{s}) + \delta)^2 - (f(\tilde{s}) - \delta)^2} = 2\sqrt{\delta f(\tilde{s})}$. Thus, $d(r,s) \leq d(a,s) \leq 2\sqrt{\delta f(\tilde{s})}$. Observe that

$$\angle rxs = \sin^{-1} \frac{d(r,s) \cdot \sin|\theta_s|}{d(r,x)} \leq \sin^{-1} \frac{2\sqrt{\delta f(\tilde{s})} \cdot |\theta_s|}{d(r,x)}.$$

Since $\delta \leq 1/(25\rho^2)$ and $|\theta_s| \leq \pi/5$, we have

$$\frac{2\sqrt{\delta f(\tilde{s})} \cdot |\theta_s|}{d(r,x)} = \frac{2\sqrt{\delta f(\tilde{s})} \cdot |\theta_s|}{f(\tilde{s}) - \delta} = \frac{2\sqrt{\delta} \cdot |\theta_s|}{\sqrt{f(\tilde{s})} - \delta/\sqrt{f(\tilde{s})}} \leq \frac{2\sqrt{\delta} \cdot |\theta_s|}{1 - \delta} < 0.06. \quad (3.15)$$

Combing (3.15) with the following fact

$$x \leq 0.6 \Rightarrow \sin^{-1} x < 1.1x, \quad (3.16)$$

we get $\angle rxs < \frac{2.2\sqrt{\delta f(\tilde{s})} \cdot |\theta_s|}{d(r,x)}$. Since $d(b,s_1) \leq d(r,s_1) = d(r,x) \cdot 2\sin\frac{\angle rxs}{2}$, we get

$$d(b,s_1) \leq d(r,x) \cdot \angle rxs \leq 2.2\sqrt{\delta f(\tilde{s})} \cdot |\theta_s|.$$

Let γ' be the acute angle between the normals at b and s_1. By Lemma 2.2.3, $\gamma' \leq 2\sin^{-1}\frac{d(b,s_1)}{(1-\alpha)f(\tilde{s})} \leq 2\sin^{-1}\frac{2.2\sqrt{\delta}\cdot|\theta_s|}{1-\alpha} \leq 2\sin^{-1}\frac{2.2\sqrt{\delta}\cdot|\theta_s|}{1-\delta}$. By (3.15) and (3.16), we conclude that $\gamma' < \frac{4.84\sqrt{\delta}\cdot|\theta_s|}{1-\delta} < 0.2|\theta_s|$. It follows that

$$\theta_s - 0.2|\theta_s| \leq \theta_s - \gamma' \leq \gamma_b \leq \theta_s + \gamma' \leq \theta_s + 0.2|\theta_s|.$$

Next, we prove the upper and lower bounds for γ_p for any point $p \in S_\alpha \cap$ *candidate*(s,θ). Let η be the acute angle between bp and the line that passes through b and is perpendicular to *candidate*(s,θ). See Figure 3.13(b). By Lemma 3.3.4(vii), the acute angle between bp and the tangent at b is at most $\sin^{-1}(0.06)$. It follows that $\eta \leq \gamma_b + \sin^{-1}(0.06) \leq \theta_s + 0.2|\theta_s| + \sin^{-1}(0.06) \leq 1.2(\pi/5) + \sin^{-1}(0.06) < 0.9$. Thus,

$$d(b,p) \leq \frac{W_s}{2\cos\eta} < 0.9W_s.$$

Note that $W_s \leq \text{radius}(\textit{coarse}(s))/3 \leq (5\rho\delta + \psi_m)f(\tilde{s})/3$, which is less than $0.02f(\tilde{s})$ by Lemma 3.3.4(i). Also, by Lemma 3.3.4(v), $f(\tilde{p}) \geq 0.9f(\tilde{s})$. It follows that

$$d(b,p) < 0.9W_s \leq 0.02f(\tilde{p}). \tag{3.17}$$

So we can invoke Lemma 2.2.3 to bound the angle γ'' between the normals at b and p:

$$\gamma'' \leq 2\sin^{-1}\frac{d(b,p)}{(1-\alpha)f(\tilde{p})} \leq 2\sin^{-1}\frac{0.9W_s}{(1-\alpha)f(\tilde{p})} \leq 2\sin^{-1}\frac{W_s}{f(\tilde{p})}.$$

By (3.17), $W_s/f(\tilde{p}) < 0.03$. So by (3.16), we get $\gamma'' \leq 2.2W_s/f(\tilde{p})$. Since $f(\tilde{p}) \geq 0.9f(\tilde{s})$, we conclude that $\gamma'' < 3W_s/f(\tilde{s})$. This implies that

$$\theta_s - 0.2|\theta_s| - 3W_s/f(\tilde{s}) \leq \gamma_b - \gamma'' \leq \gamma_p \leq \gamma_b + \gamma'' \leq \theta_s + 0.2|\theta_s| + 3W_s/f(\tilde{s}).$$

Finally, we have proved the lemma under the conditions that $\max\{2\sqrt{\rho}\delta, \psi_h\sqrt{f(\tilde{s})}\} \leq \text{radius}(\textit{coarse}(s)) \leq 5\rho\delta + \psi_m\sqrt{f(\tilde{s})}$ and $\text{radius}(\textit{initial}(s)) \leq \psi_m\sqrt{f(\tilde{s})}$. These conditions hold with probabilities at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$ by Lemmas 3.3.1, 3.3.2, and 3.3.3. So the lemma follows. □

Proof of Lemma 3.4.2

Let ϕ be the acute angle between uv and the tangent to S_α at u. Let η be the acute angle between uv and the direction of $candidate(s,\theta)$. By Lemma 3.3.4(vii), $\phi \leq \sin^{-1}(0.06)$. So $\eta \geq \pi/2 - \gamma_u - \phi \geq \pi/2 - \gamma_u - \sin^{-1}(0.06)$. By Lemma 3.4.1(i), (ii), and (iv), $\eta \geq \pi/2 - 1.2(\pi/5) - 3(0.1) - \sin^{-1}(0.06) > 0.4$. Thus, $d(u,v) \leq \frac{\text{width}(H)}{\sin\eta} \leq \frac{\text{width}(H)}{\sin(0.4)} < 3\,\text{width}(H)$. This proves (i).

Consider (ii). Note that $W_s \leq \text{radius}(coarse(s))/3 \leq (5\rho\delta + \psi_m)f(\tilde{s})/3$. So by (i), $d(u,v) \leq 3W_s \leq (5\rho\delta + \psi_m)f(\tilde{s})$. By Lemma 3.3.4(i) and (v), $5\rho\delta + \psi_m \leq 0.05$ and $f(\tilde{u}) \geq 0.9f(\tilde{s})$. It follows that

$$d(u,v) < 0.06f(\tilde{u}). \tag{3.18}$$

Thus, we can invoke Lemma 2.2.3 to bound the angle ξ between the normals at u and v:

$$\xi \leq 2\sin^{-1}\frac{d(u,v)}{(1-\alpha)f(\tilde{u})} \leq 2\sin^{-1}\frac{3\,\text{width}(H)}{0.9(1-\alpha)f(\tilde{s})} < 2\sin^{-1}\frac{4\,\text{width}(H)}{f(\tilde{s})}.$$

Since $4\,\text{width}(H)/f(\tilde{s}) \leq 4W_s/f(\tilde{s})$ which is at most 0.4 by Lemma 3.4.1(i), we can apply (3.16) to conclude that $\xi < 9\,\text{width}(H)/f(\tilde{s}) \leq 9\,\text{width}(H)$. This proves (ii).

Finally, by (3.18), we can invoke Lemma 2.2.2(ii) to bound the acute angle between uv and the tangent at u. This angle is at most $\sin^{-1}\frac{d(u,v)}{2(1-\alpha)f(\tilde{u})}$ which is less than $\xi/2$. □

Proof of Lemma 3.5.1

We prove the lemma by assuming that Lemma 3.3.1, 3.3.2, and 3.3.3 hold deterministically. The probability bound then follows from the probability bounds in these lemmas. For $i = p$ or q, let $R_i = \text{radius}(coarse(i))$ and let $r_i = \text{radius}(initial(i))$. The Lipschitz condition implies that $f(\tilde{p})/2 \leq f(\tilde{q}) \leq 3f(\tilde{p})/2$. Let h and m be the constants in Lemma 3.3.1.

Suppose that $W_p = \sqrt{r_p}$. By Lemma 3.3.1, we have

$$W_p = \sqrt{r_p} \geq \sqrt{\frac{\lambda_h\sqrt{f(\tilde{p})}}{3}} = \sqrt{\frac{h\lambda_m\sqrt{f(\tilde{p})}}{3m}}.$$

Note that $W_q \leq \sqrt{r_q}$ and $r_q \leq \sqrt{14\lambda_m f(\tilde{q})}$ by Lemma 3.3.1. So we get

$$W_p \geq \sqrt{\frac{h\sqrt{f(\tilde{p})}}{42mf(\tilde{q})}} \cdot r_q \geq \sqrt{\frac{h}{63m\sqrt{f(\tilde{p})}}} \cdot W_q^2 \geq \sqrt{\frac{h}{63m}} \cdot \frac{W_q^2}{f(\tilde{p})}.$$

Suppose that $W_p = R_p/3$. First, since $R_p \geq 2\sqrt{\rho}\delta$ by Lemma 3.3.3, we get $\rho\delta \leq 3\sqrt{\rho}W_p/2$. Second, $W_p = R_p/3 \geq r_p/3$ which is at least $\lambda_h\sqrt{f(\tilde{p})}/9$ by Lemma 3.3.1. So we get $\sqrt{\lambda_m f(\tilde{p})} = \sqrt{m\lambda_h f(\tilde{p})/h} \leq 3\sqrt{mW_p/h} \cdot f(\tilde{p})^{1/4} \leq 3\sqrt{mW_p/h} \cdot f(\tilde{p})$. Finally, since $W_q \leq R_q/3$, by Lemma 3.3.2, we get

$$\begin{aligned} W_q &\leq \frac{5\rho\delta}{3} + \frac{\sqrt{14\lambda_m f(\tilde{q})}}{3} \\ &\leq \frac{5\rho\delta}{3} + \sqrt{\frac{7\lambda_m f(\tilde{p})}{3}} \\ &\leq \frac{5\sqrt{\rho}W_p}{2} + \sqrt{\frac{21mW_p}{h}} \cdot f(\tilde{p}). \end{aligned}$$

□

Proof of Lemma 3.5.3

We prove the lemma by assuming that Lemmas 3.4.4 and 3.5.1 hold deterministically. The probability bound then follows from the probability bounds in these lemmas.

We translate x^*y^* to align y^* with $\tilde{y}$. Let z denote the point $x^* + \tilde{y} - y^*$. Let $k = 138\delta + 3$. By triangle inequality and Lemma 3.4.4, $d(\tilde{x}, z) \leq d(x^*, \tilde{x}) + d(y^*, \tilde{y}) \leq kW_x + kW_y$. Since $d(\tilde{x}, \tilde{y}) \leq f(\tilde{y})/2$, by Lemma 3.5.1, $W_x \leq \mu_1 f(\tilde{y})\sqrt{W_y}$. So $d(\tilde{x}, z) \leq k\mu_1 f(\tilde{y})\sqrt{W_y} + kW_y$, which is smaller than $W_y^{1/3} \leq d(x^*, y^*)$ for sufficiently large n. Thus, $\tilde{x}z$ is not the longest side of the triangle $\tilde{x}\tilde{y}z$. It follows that $\angle\tilde{x}\tilde{y}z$ is acute. Since $d(\tilde{x}, z)$ is an upper bound on the height of z from $\tilde{x}\tilde{y}$, we have $\angle\tilde{x}\tilde{y}z \leq \sin^{-1}\frac{d(\tilde{x},z)}{d(\tilde{y},z)} = \sin^{-1}\frac{d(\tilde{x},z)}{d(x^*,y^*)} \leq \sin^{-1}(k\mu_1 f(\tilde{y})W_y^{1/6} + kW_y^{2/3})$. We conclude that $\angle\tilde{x}\tilde{y}z$ is $O(f(\tilde{y})W_y^{1/6})$ as n tends to ∞. □

Proof of Lemma 3.5.4

We first show that $d(\tilde{x},\tilde{z}) \leq \min\{f(\tilde{x})/4, f(\tilde{z})/4\}$. Assume that $d(\tilde{x},\tilde{z}) \leq f(\tilde{x})/5$. By the Lipschitz condition, we have $f(\tilde{z}) \geq 4f(\tilde{x})/5$. Therefore, $d(\tilde{x},\tilde{z}) \leq f(\tilde{x})/5 \leq f(\tilde{z})/4$.

Let D be the disk centered at $\tilde{x}$ with radius $f(\tilde{x})/4$. Observe that $S(\tilde{x},\tilde{z})$ lies completely inside D. Otherwise, the medial axis of S intersects the interior of D which implies that $f(\tilde{x}) \leq f(\tilde{x})/4$, a contradiction. So $d(\tilde{x},\tilde{y}) \leq f(\tilde{x})/4$. The Lipschitz condition implies that $f(\tilde{y}) \geq 3f(\tilde{x})/4$.

We claim that the angle $\angle\tilde{x}\tilde{y}\tilde{z}$ is obtuse. The line segments $\tilde{x}\tilde{y}$ and $\tilde{y}\tilde{z}$ are parallel to the tangents at some points on $S(\tilde{x},\tilde{y})$ and $S(\tilde{y},\tilde{z})$, respectively. Lemma 2.2.3 implies that $\angle\tilde{x}\tilde{y}\tilde{z} \geq \pi - 4\sin^{-1}\frac{\text{radius}(D)}{f(\tilde{x})} = \pi - 4\sin^{-1}(1/4) > \pi/2$.

Since $d(\tilde{x},\tilde{y}) \leq f(\tilde{x})/4 \leq f(\tilde{y})/3$, by Lemma 3.5.3, the angle between x^*y^* and $\tilde{x}\tilde{y}$ is negligible with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$ as n tends to ∞. A symmetric argument shows that the angle between y^*z^* and $\tilde{y}\tilde{z}$ is negligible with probability at least $1 - O(n^{-\Omega(\ln^\omega n/f_{\max})})$ as n tends to ∞. Thus, $\angle x^*y^*z^*$ converges to $\angle\tilde{x}\tilde{y}\tilde{z}$ which is obtuse. □

Proof of Lemma 3.5.5

Note that p^* and q^* are adjacent and they are selected by the algorithm. Let $k = 138\delta + 3$. Let D_p be the disk centered at p^* with radius $(1 + k\mu_1 f(\tilde{p}))W_p^{1/3}$. Let D_q be the disk centered at q^* with radius $(1 + k\mu_1 f(\tilde{q}))W_q^{1/3}$. By Lemma 3.4.4, $d(\tilde{p},p^*) \leq kW_p$ which is less than $W_p^{1/3}$ for sufficiently large n. So $\tilde{p}$ lies inside D_p. Similarly, $\tilde{q}$ lies inside D_q.

If D_p intersects D_q, then $d(p^*,q^*) \leq (1 + \mu_1 f(\tilde{p}))W_p^{1/3} + (1 + \mu_1 f(\tilde{q}))W_q^{1/3}$ and we are done. Suppose that D_p does not intersect D_q. We claim that $S(\tilde{p},\tilde{q}) \cap D_p$ is connected. Otherwise, the medial axis of S intersects the interior of D_p which implies that $f(\tilde{p}) \leq \text{radius}(D_p)$ which is less than $f(\tilde{p})$ for sufficiently large n, a contradiction. Similarly, $S(\tilde{p},\tilde{q}) \cap D_q$ is connected. It follows that $S(\tilde{p},\tilde{q}) - (D_p \cup D_q)$ is also connected. There are two cases.

Case 1: $S(\tilde{p},\tilde{q}) - (D_p \cup D_q)$ does not contain $\tilde{u}$ for any sample u. Let y be the endpoint of $S(\tilde{p},\tilde{q}) - (D_p \cup D_q)$ that lies on D_p. Let h be the constant in Lemma 3.3.1. Take a λ_h-partition such that y is the first cut-point. Since $S(\tilde{p},\tilde{q}) - (D_p \cup D_q)$ does not contain $\tilde{u}$ for any sample u, by Lemma 2.3.6(i), $S(\tilde{p},\tilde{q}) - (D_p \cup D_q)$ does not contain $S(y, c_1)$, where c_1 is the second cut-point, with probability at least $1 - O(n^{-\Omega(\ln^\omega n)})$. It follows that

$$|S(\tilde{p},\tilde{q}) - (D_p \cup D_q)| < \lambda_h^2 f(y). \tag{3.19}$$

Since $d(\tilde{p}, y) \leq 2\,\mathrm{radius}(D_p) = 2(1 + k\mu_1 f(\tilde{p}))W_p^{1/3}$, $d(\tilde{p}, y) \leq f(\tilde{p})/2$ for sufficiently large n. Thus, $f(y) \leq 3f(\tilde{p})/2$, so $\lambda_h^2 f(y) < 3\lambda_h^2 f(\tilde{p})/2$. Since $W_p \geq \mathrm{radius}(\mathit{initial}(p))/3$ which is at least $\lambda_h\sqrt{f(\tilde{p})}/9$ by Lemma 3.3.1, we have $\lambda_h^2 f(\tilde{y}) \leq 243W_p^2/2$. Substituting into (3.19), we get

$$|S(\tilde{p},\tilde{q})| \leq 243W_p^2/2 + 2\,\mathrm{radius}(D_p) + 2\,\mathrm{radius}(D_q).$$

By Lemma 3.4.4, $d(\tilde{p}, p^*) \leq kW_p$ and $d(\tilde{q}, q^*) \leq kW_q$. We conclude that $d(p^*, q^*) \leq d(\tilde{p}, p^*) + |S(\tilde{p},\tilde{q})| + d(\tilde{q}, q^*) \leq \mu_2 f(\tilde{p})W_p^{1/3} + \mu_2 f(\tilde{q})W_q^{1/3}$ for some constant $\mu_2 > 0$.

Case 2: $S(\tilde{p},\tilde{q}) - (D_p \cup D_q)$ contains $\tilde{u}$ for some sample u. We show that this case is impossible if Lemmas 3.5.1 and 3.5.4 hold deterministically. It follows that case 2 occurs with probability at most $O(n^{-\Omega(\ln^\omega n/f_{\max})})$. We first claim that $d(p^*, u^*) > W_p^{1/3}$. If not, Lemma 3.5.1 implies that $W_u \leq \mu_1 f(\tilde{p})\sqrt{W_p}$ for sufficiently large n. But then $d(p^*, \tilde{u}) \leq d(p^*, u^*) + d(\tilde{u}, u^*) \leq W_p^{1/3} + kW_u \leq W_p^{1/3} + k\mu_1 f(\tilde{p})\sqrt{W_p}$. This is a contradiction as $\tilde{u}$ lies outside D_p. Similarly, $d(q^*, u^*) > W_q^{1/3}$. So u^* is not eliminated by the selection of p^* and q^*.

Next, take any selected center point z^* different from p^* and q^* such that $\tilde{q} \in S(\tilde{u}, \tilde{z})$. We show that u^* is not eliminated by the selection of z^*. Assume to the contrary that this is false. So $d(u^*, z^*) \leq W_z^{1/3}$. By Lemma 3.5.1, $W_u \leq \mu_1 f(\tilde{z})\sqrt{W_z}$ for sufficiently large n. Let $k' = 1 + k + k\mu_1$. Then $d(\tilde{u}, \tilde{z}) \leq d(u^*, z^*) + d(z^*, \tilde{z}) + d(u^*, \tilde{u}) \leq W_z^{1/3} + kW_z + kW_u \leq W_z^{1/3} + kW_z + k\mu_1 f(\tilde{z})\sqrt{W_z} \leq k' f(\tilde{z})W_z^{1/3}$. For sufficiently large n, $k' f(\tilde{z})W_z^{1/3} \leq f(\tilde{z})/5$.

By Lemma 3.5.4, the angle $\angle u^*q^*z^*$ is obtuse. It follows that $d(q^*, z^*) < d(u^*, z^*) \leq W_z^{1/3}$, contradicting Lemma 3.5.2.

Symmetrically, we can show that u^* is not eliminated by any selected center point z^* different from p^* and q^* such that $\tilde{p} \in S(\tilde{z}, \tilde{u})$. In all, our algorithm should select another center point u^* such that $\tilde{u} \in S(\tilde{p}, \tilde{q}) - (D_p \cup D_q)$. This contradicts the assumption that p^* and q^* are adjacent in G.

□

CHAPTER 4

DETERMINISTIC NOISE MODEL FOR SURFACES

In this chapter, we extend and improve the noise model in Chapter 2 to surfaces in three dimensions. We use S to denote the smooth closed surface and P to denote the set of noisy sample points from S. The *noise amplitude* of P is the maximum of the distances from P to the surface S. We denote the noise amplitude of P to S by δ.

First, we introduce some basic notations and some basic geometric lemmas in Section 4.1. In Section 4.2, we present a deterministic noise model for surfaces. In Section 4.3, we justify this deterministic model using a variant of the probabilistic noise model defined in Chapter 2. Note that the proofs of some technical lemmas are given in the appendix of this chapter.

4.1 Preliminaries

Recall that the *medial axis* of S is the set of centers of empty balls that touch S at more than one point, and for any pint $x \in S$, the *local feature size* $f(x)$ is the minimum distance from x to the medial axis of S. We assume that $\min_{x \in S} f(x) = 1$ for convenience. So for any $x \in S$, $f(x) \geq 1$. We suppose the noise amplitude $\delta < 1$ as we only consider noise of first type.

For any $\alpha \in [-\delta, \delta]$, S_α^+ (resp., S_α^-) is the collection of points in R^3 that lie outside (resp., inside) S and their distances from S are exactly α. We use S_α when it is unimportant to distinguish between inside and outside. Note that S_α and S have the same medial axis. We use $\mathbf{S}$ to denote the set of points sandwiched between S_δ^+ and S_δ^-.

Given two objects A and B, we use $d(A, B)$ to denote the minimum distance between A and B. Given two coplanar vectors/lines in 3D, the plane *spanned by*

them is the plane that contains them. Given two vectors, lines, or planes A and B in 3D, we use $\angle(A, B)$ to denote the smallest angle between them. If A and B are disjoint, we translate one of them to measure $\angle(A, B)$.

For each point $p \in S_\alpha$, we use n_p to denote the *outward unit normal* of S_α at p, and T_p to denote the *tangent plane* of S_α at p. We use $\tilde{p}$ to denote the point on S such that $d(p, \tilde{p}) = \alpha$. Note that n_p and $n_{\tilde{p}}$ have the same support line, and T_p and $T_{\tilde{p}}$ are parallel. We use $B(p, d)$ to denote a ball centered at p with radius d.

We state three geometric lemmas that will be useful later. Their proofs can be found in the appendix.

Lemma 4.1.1 *Any point p on S_α has two tangent balls with radii $f(\tilde{p}) - \alpha$ whose interior do not intersect S_α.*

The next lemma follows from results in [2, 29]. It basically says that the neighborhood of any point on S_α is fairly flat, and the normal varies slowly.

Lemma 4.1.2 *Assume that $\delta \leq 1/6$. Let $p, q \in S_\alpha$ be two points.*

(i) If $d(p, q) = \lambda(f(\tilde{p}) - \delta)$ for some $\lambda < 1$, then $d(q, T_p) \leq d(p, q)^2$ and $\angle(pq, T_p) \leq \arcsin(d(p, q))$.

(ii) If $d(p, q) = \lambda \min\{f(\tilde{p}) - \delta, f(\tilde{q}) - \delta\}$ for some $\lambda \leq 1/5$, then $\angle(n_p, n_q) \leq 3d(p, q)$.

The following lemma shows that the distance between nearby points on S are preserved in S_α up to a constant factor.

Lemma 4.1.3 *Assume $\delta \leq 1/10$. Let $p, q \in S_\alpha$ be two points such that $d(\tilde{p}, \tilde{q}) = d$ for some $d \leq 1/5$. Then $d/2 \leq d(p, q) \leq 3d/2$.*

4.2 Sampling and noise model

Let ω be a fixed positive constant (we think of it as very small). We assume that the set P of noisy samples satisfies the following conditions.

There exists parameters $0 < \epsilon_2 \leq \epsilon \leq \epsilon_1 < 1$ depending on n such that

(i) $\epsilon_1 \to 0$ as $n \to \infty$,

(ii) $\epsilon_1 = O(\epsilon)$, $\epsilon_2 = \Omega(\epsilon^c)$ for some constant $c \geq 1$.

and for any point $p \in \mathbf{S}$, we have

(i) $|B(p, \epsilon) \cap P| \geq 1$,

(ii) $|B(p, \epsilon_1) \cap P| \geq 2 \ln^{\omega} n$,

(iii) $|B(p, \epsilon_2) \cap P| \leq \frac{1}{2} \ln^{\omega} n$.

In Theorem 4.3.1, specifically, we can prove that for $\epsilon = \Theta(\sqrt[3]{\frac{\ln^{\omega} n}{n}})$, $\epsilon_1 = \Theta(\sqrt[3]{\frac{\ln^{\omega} n}{n}})$ and $\epsilon_2 = \Theta(\sqrt{\frac{\ln^{\omega} n}{n}})$, our deterministic model is satisfied with high probability by a set of sample points generated according to a uniform distribution. In a sense, the parameters $\epsilon, \epsilon_1, \epsilon_2$ measure the "sampling density". It would be nice that the three parameters collapse to a single one (with some appropriate adjustment in the conditions), but our analysis in the justification is not strong enough to do so. It is also more natural that $\epsilon_2 = \Omega(\epsilon)$, but we can only justify our model for $c \geq 3/2$. Therefore, we will assume throughout this paper that $\epsilon_2 = \Omega(\epsilon^{3/2})$. We remark that our algorithm does not know the values of ϵ, ϵ_1, ϵ_2, and the noise amplitude δ.

4.3 Justification

Consider generating the noisy samples using the following random process. First, points are drawn from S according to a uniform distribution. That is, the probability of drawing a point from a region A is equal to $\frac{\text{area}(A)}{\text{area}(S)}$. Second, each point x drawn is perturbed uniformly within a line segment with width 2δ, centered at x, and aligned with n_x. Moreover, the distribution of each sample point is independently identical. A similar model was used in the curve case in Chapter 2. Our goal is to prove that the samples generated by the above random process satisfy our deterministic model with probability at least $1 - \exp(-\Omega(\ln^{\omega} n))$.

For any constant $k > 0$, let λ_k denote $\sqrt[3]{\frac{k \ln^\omega n}{n}}$. Note that $0 < \lambda_k < 1$ for sufficiently large n. Let s be a point on S_α. Let $R_1 \subseteq S_\alpha$ denote the patch $\{p \in S_\alpha : d(p,s) \leq \lambda_k\}$. Let $R_2 \subseteq S_\alpha$ denote the patch $\{p \in S_\alpha : d(p,s) \leq \lambda_k^{3/2}\}$. The *$\lambda_k$-cell at s* is the union of line segments with midpoints on R_1, normal to R_1, and with length $2\lambda_k\delta$. Note that the λ_k-cell may not lie completely inside $\mathbf{S}$, but at least half of it on one side of R_1 does. The *λ_k-rod at s* is the union of line segments with length 2δ, inside $\mathbf{S}$, and normal to some point in R_2. The following two lemmas bound the sizes of a λ_k-cell and a λ-rod.

Lemma 4.3.1 *The λ_k-cell at a point s lies inside $B(s, 2\lambda_k)$.*

Proof. Assume that $s \in S_\alpha$. Let C denote the λ_k-cell at s. Let p be any point in C. Let q be the projection of p along the normal to S_α, i.e., $\tilde{p} = \tilde{q}$. Then $d(p,q) \leq \lambda_k\delta$ and $d(q,s) \leq \lambda_k$. Thus $d(p,s) \leq d(p,q) + d(q,s) \leq \lambda_k\delta + \lambda_k \leq 2\lambda_k$. ◻

Lemma 4.3.2 *The λ_k-rod at a point s contains $B(s, \lambda_k^{3/2}/4) \cap \mathbf{S}$.*

Proof. Let C denote the λ_k-rod at s. Let p be a point in $B(s, \lambda_k^{3/2}/4) \cap \mathbf{S}$. Assume that $p \in S_\alpha$. Let s' denote the projection of s onto S_α along the normal direction. Clearly $d(s',s) \leq d(s,p) \leq \lambda_k^{3/2}/4$. Thus $d(s',p) \leq d(s',s) + d(s,p) \leq \lambda_k^{3/2}/2$. Assume that $s \in S_\beta$. Let p' denote the projection of p onto S_β along the normal direction. By Lemma 4.1.3, $d(s,p') \leq 2d(s',p) \leq \lambda_k^{3/2}$. Thus $p' \in C \cap S_\beta$, which implies that $p \in C \cap S_\alpha$. ◻

We give a highlight of our proof. We first upper and lower bound the probabilities of a sample appearing in the λ_k-cell and λ_k-rod at s. Based on these probability bounds, we can show that the number of samples inside the λ_k-cell at s is essentially $\Theta(\ln^\omega n)$ with high probability. The same holds for the λ_k-rod at s. By setting ϵ_1 and ϵ_2 appropriately, we then show that the λ_k-cell lies inside the ball $B(s, \epsilon_1)$ and the ball $B(s, \epsilon_2)$ lies inside the λ_k-rod. Hence the conditions in our deterministic noise model are satisfied. We give the details of the proof below.

The next technical lemma follows from the fact that a small neighborhood of any point on S_α is fairly flat.

Lemma 4.3.3 *Assume that $r \leq (1-\delta)/20$. Let p be a point on S_α. Let Y be the infinite cylinder with radius r and axis aligned with n_p. Then the area of the connected component in $Y \cap S_\alpha$ containing p is $\Theta(r^2)$.*

Next, we bound the probability of a sample appearing in the λ_k-cell and λ_k-rod.

Lemma 4.3.4 *Let C be the λ_k-cell or λ_k-rod at a point $s \in \mathbf{S}$. Assume that n is so large that $\lambda_k \leq 1/4$. Then for any sample p, $\kappa_2 \lambda_k^3 \leq \Pr(p \in C) \leq \kappa_1 \lambda_k^3$ for some constants κ_1 and κ_2.*

Proof. Consider the case that C is a λ_k-cell. Assume that s lies on S_α. Let $R \subseteq S$ denote the patch $\{\tilde{x} : x \in C \cap S_\alpha\}$. Project R orthogonally onto the region R' on the tangent plane $T_{\tilde{s}}$ at $\tilde{s}$. Take a point q in the boundary of $C \cap S_\alpha$. Let q' denote the orthogonal projection of $\tilde{q}$ onto $T_{\tilde{s}}$. By definition, $d(s,q) = \lambda_k$. Then Lemma 4.1.3 implies that $\lambda_k/2 \leq d(\tilde{s},\tilde{q}) \leq 2\lambda_k$.

Clearly, $d(\tilde{s},q') \leq d(\tilde{s},\tilde{q}) \leq 2\lambda_k$. Thus R' lies within the disk on $T_{\tilde{s}}$ centered at $\tilde{s}$ with radius $2\lambda_k$. Then by Lemma 4.3.3, $\text{area}(R) = O(\lambda_k^2)$. On the other hand, $d(\tilde{s},q') = d(\tilde{s},\tilde{q}) \cos\angle(\tilde{q}\tilde{s}, T_{\tilde{s}}) \geq \lambda_k/2 \cdot \cos\angle(\tilde{q}\tilde{s}, T_{\tilde{s}})$. By Lemma 4.1.2(i), $\angle(\tilde{q}\tilde{s}, T_{\tilde{s}}) \leq \arcsin(d(\tilde{s},\tilde{q})) \leq 2d(\tilde{s},\tilde{q}) \leq 4\lambda_k$. Since $\lambda_k \leq 1/4$ by assumption, $\angle(\tilde{q}\tilde{s}, T_{\tilde{s}}) \leq 1$. So $\cos\angle(\tilde{q}\tilde{s}, T_{\tilde{s}}) \geq \cos(1) \geq 1/2$. This implies that $d(\tilde{s},q') \geq \lambda_k/4$. Thus R' contains a disk on $T_{\tilde{s}}$ centered at $\tilde{s}$ with radius $\lambda_k/4$. Combining with the previous upper bound, we conclude that $\text{area}(R) = \Theta(\lambda_k^2)$.

The probability of drawing a point from R is thus $\Theta(\lambda_k^2)/\text{area}(S)$. The probability that the perturbation throws the point drawn from R into C is at least $\frac{\lambda_k \delta}{2\delta} = \frac{\lambda_k}{2}$, and at most $\frac{2\lambda_k \delta}{2\delta} = \lambda_k$. It follows that for any sample p, $\Pr(p \in C) = \Theta(\lambda_k^3)$, assuming that $\text{area}(S)$ is an intrinsic constant for S.

Now consider the case that C is a λ_k-rod. We also let $R \subseteq S$ denote the patch $\{\tilde{x} : x \in C \cap S_\alpha\}$. By the definition of a λ_k-rod and an argument similar to the one in the above, we can show that $\text{area}(R) = \Theta(\lambda_k^3)$. Then the probability of a

sample falling inside C is same as the probability of drawing a point from R, which is area(R)/area$(S) = \Theta(\lambda_k^3)$. □

Next, we bound the number of samples inside the λ_k-cell and λ_k-rod in Lemma 4.3.6. The following Chernoff bound [32] stated in Lemma 4.3.5 will be needed for doing that.

Lemma 4.3.5 *Let the random variables $X_1, X_2, \ldots, X_n$ be independent, with $0 \leq X_i \leq 1$ for each i. Let $S_n = \sum_{i=1}^{n} X_i$, and let $E(S_n)$ be the expected value of S_n. Then for any $\sigma > 0$, $\Pr(S_n \leq (1-\sigma)E(S_n)) \leq \exp(-\frac{\sigma^2 E(S_n)}{2})$, and $\Pr(S_n \geq (1+\sigma)E(S_n)) \leq \exp(-\frac{\sigma^2 E(S_n)}{2(1+\sigma/3)})$.*

Lemma 4.3.6 *Assume that n is so large that $\lambda_k \leq 1/4$. Let C be a λ_k-rod or λ_k-cell at a point $s \in \mathbf{S}$. Let κ_1 and κ_2 be the constants in Lemma 4.3.4.*

(i) C is non-empty with probability at least $1 - \exp(-\Omega(\ln^{\omega} n))$.

(ii) For any constant $\kappa > \kappa_1 k$, the number of samples in C is at most $\kappa \ln^{\omega} n$ with probability at least $1 - \exp(-\Omega(\ln^{\omega} n))$.

(iii) For any constant $\kappa < \kappa_2 k$, the number of samples in C is at least $\kappa \ln^{\omega} n$ with probability at least $1 - \exp(-\Omega(\ln^{\omega} n))$.

Proof. Let $X_i (i = 1, \ldots, n)$ be a random binomial variable taking value 1 if the sample point s_i is inside C, and value 0 otherwise. Let $S_n = \sum_{i=1}^{n} X_i$. Then $E(S_n) = \sum_{i=1}^{n} E(X_i) = n \cdot \Pr(s_i \in C)$. This implies that

$$E(S_n) \leq \kappa_1 n \lambda_k^3 = \kappa_1 k \ln^{\omega} n, \text{ and}$$
$$E(S_n) \geq \kappa_2 n \lambda_k^3 = \kappa_2 k \ln^{\omega} n$$

By Lemma 4.3.5,

$$\begin{aligned} \Pr(S_n \leq 0) &= \Pr(S_n \leq (1-1)E(S_n)) \\ &\leq \exp(-\frac{E(S_n)}{2}) \\ &\leq \exp(-\Omega(\ln^{\omega} n)). \end{aligned}$$

Consider (ii). Let $\sigma = \frac{\kappa}{\kappa_1 k} - 1 > 0$. We have

$$\kappa \ln^\omega n = \kappa_1 n \lambda_k^3 (1+\sigma) \geq (1+\sigma) E(S_n).$$

By Lemma 4.3.5,

$$\begin{aligned} \Pr(S_n > \kappa \ln^\omega n) &\leq \Pr(S_n > (1+\sigma)E(S_n)) \\ &\leq \exp(-\frac{\sigma^2 E(S_n)}{2+2\sigma/3}) \\ &= \exp(-\Omega(\ln^\omega n)). \end{aligned}$$

Consider (iii). Let $\sigma = 1 - \frac{\kappa}{\kappa_2 k} > 0$. We have

$$\kappa \ln^\omega n = \kappa_2 n \lambda_k^3 (1-\sigma) \leq (1-\sigma) E(S_n).$$

By Lemma 4.3.5,

$$\begin{aligned} \Pr(S_n < \kappa \ln^\omega n) &\leq \Pr(S_n < (1-\sigma)E(S_n)) \\ &\leq \exp(-\frac{\sigma^2 E(S_n)}{2}) \\ &= \exp(-\Omega(\ln^\omega n)). \end{aligned}$$

□

We are ready to prove that the random process produces samples that satisfy our deterministic model with high probability.

Theorem 4.3.1 *Let X be n points on S that are drawn uniformly. For each $x \in X$, define $p_x = x + \alpha_x n_x$, where α_x is drawn uniformly and independently from $[-\delta, \delta]$. There exists $\epsilon, \epsilon_1, \epsilon_2$ such that $\{p_x : x \in X\}$ satisfy our deterministic model with probability at least $1 - \exp(-\Omega(\ln^\omega n))$.*

Proof. Choose two constants a and b such that $2 < \kappa_2 a$ and $\frac{1}{2} > \kappa_1 b$. We prove the theorem for $\epsilon = 2\lambda_1$, $\epsilon_1 = 2\lambda_a$, and $\epsilon_2 = \lambda_b^{3/2}/4$. First of all, it is clear that $\epsilon_1 = O(\epsilon)$ and $\epsilon_2 = \Omega(\epsilon^{3/2})$. For convenience, we use "with high probability" to mean "with probability at least $1 - \exp(-\Omega(\ln^\omega n))$".

Let p be a point in $\mathbf{S}$. Let C be the λ_1-cell at p. By Lemma 4.3.6(i), C contains a sample with high probability. By Lemma 4.3.1, C lies inside the ball $B(p, 2\lambda_1) = B(p, \epsilon)$. Thus $|B(p, \epsilon) \cap P| \geq 1$ with high probability.

Let C_1 be the λ_a-cell at p. By Lemma 4.3.6(iii), C_1 contains at least $2 \ln^{\omega} n$ samples with high probability. By Lemma 4.3.1, C_1 lies inside the ball $B(p, 2\lambda_a) = B(p, \epsilon_1)$. Thus $|B(p, \epsilon_1) \cap P| \geq 2 \ln^{\omega} n$ with high probability.

Let C_2 be the λ_b-rod at p. By Lemma 4.3.6(ii), C_2 contains at most $\frac{1}{2} \ln^{\omega} n$ samples with high probability. By Lemma 4.3.2, the ball $B(p, \lambda_b^{3/2}/4) \cap \mathbf{S} = B(p, \epsilon_2) \cap \mathbf{S}$ lies inside C_2. Thus $|B(p, \epsilon_2) \cap P| \leq \frac{1}{2} \ln^{\omega} n$ with high probability. □

4.4 Appendix

Proof of Lemma 4.1.1

Let M_α be the medial ball of S_α touching a point $p \in S_\alpha$. By the definition of S_α, there is a medial ball M of S touching $\tilde{p}$ such that M and M_α have the same center and $\text{radius}(M_\alpha) = \text{radius}(M) - \alpha \geq f(\tilde{p}) - \alpha$. Let B be a ball of radius $f(\tilde{p}) - \alpha$ that touches S_α at p. If S_α intersects the interior of B, the medial axis of S_α intersects the interior of B. But then $\text{radius}(M_\alpha) < \text{radius}(B) = f(\tilde{p}) - \alpha$, a contradiction. □

The next lemma consists of a lemma by Giesen and Wagner [29] and a lemma by Amenta and Bern [2].

Lemma 4.4.1 *Let $p, q \in S$ be two points.*

(i) If $d(p,q) = \lambda f(p)$ for some $\lambda < 1$, then $d(q, T_p) \leq \frac{\lambda^2}{2} f(p)$ and $\angle(pq, T_p) \leq \arcsin(\frac{\lambda}{2})$.

(ii) If $d(p,q) = \lambda \min\{f(p), f(q)\}$ for some $\lambda \leq 1/3$, then $\angle(n_p, n_q) \leq \frac{\lambda}{1-3\lambda}$.

We are ready to prove Lemmas 4.1.2 and 4.1.3.

Proof of Lemma 4.1.2

Let f_α denote the local feature size function for any point on S_α. As S and S_α share the same medial axis, $f_\alpha(p) \geq f(\tilde{p}) - \delta \geq 1/2$ for any point $p \in S_\alpha$.

Consider (i). We have $d(p,q) = \beta f_\alpha(p)$ for some $\beta \leq \lambda$. Lemma 4.4.1(i) implies that

$$\begin{aligned} d(q, T_p) &\leq \beta^2 f_\alpha(p)/2 \\ &= d(p,q)^2/(2f_\alpha(p)) \\ &\leq d(p,q)^2. \end{aligned}$$

And $\angle(n_p, n_q) \leq \arcsin(\frac{\beta}{2}) \leq \arcsin(\frac{d(p,q)}{2f_\alpha(p)})$, which is at most $\arcsin(d(p,q))$.

Consider (ii). Let g denote $\min\{f(\tilde{p}) - \delta, f(\tilde{q}) - \delta\}$. We then have

$$d(p,q) = \beta \min\{f_\alpha(p), f_\alpha(q)\}$$

for some $\beta \leq \lambda \leq 1/5$. Lemma 4.4.1(ii) implies that

$$\begin{aligned} \angle(n_p, n_q) &\leq \beta/(1-3\beta) \\ &\leq d(p,q)/((1-3\beta)g). \end{aligned}$$

Note that $g \geq 1 - \delta \geq 5/6$ and $1 - 3\beta \geq 2/5$. So $(1-3\beta)g \geq 1/3$, which implies that $\angle(n_p, n_q) \leq 3d(p,q)$. □

Proof of Lemma 4.1.3

Let r be the point $\tilde{p} + (q - \tilde{q})$. W.l.o.g., assume that $\angle\tilde{p}pr \leq \angle\tilde{p}rp$. By Lemma 4.1.2, $\angle p\tilde{p}r \leq 3d$. Therefore, $\angle\tilde{p}rp \geq \pi/2 - 3d/2$. By sine law, $d(p,r) = \frac{d(p,\tilde{p}) \cdot \sin\angle p\tilde{p}r}{\sin\angle\tilde{p}rp} \leq \frac{3\delta d}{\cos(3d/2)}$. Since $\delta \leq 1/10$ and $\cos(3d/2) \geq \cos(3/10) > 0.9$, we have $d(p,r) \leq d/2$. By triangle inequality,

$$\begin{aligned} d(p,q) &\leq d(q,r) + d(p,r) \\ &= d(\tilde{p},\tilde{q}) + d(p,r) \\ &\leq 3d/2. \end{aligned}$$

Similarly, $d(p,q) \geq d(q,r) - d(p,r) \geq d/2$. □

We need the Lemmas 4.4.2, 4.4.3, and 4.4.4, to prove Lemma 4.3.3.

Lemma 4.4.2 *Let $p \in S_\alpha$ be a point. Let k and d be constants such that $k \geq 2$ and $kd \leq 1/20$. Let ℓ be a line such that $d(p,\ell) \leq d$ and $\angle(n_p, \ell) \leq \pi/4$. Then ℓ intersects $S_\alpha \cap B(p, kd)$ at exactly one point.*

Proof. Let $C = S_\alpha \cap B(p, kd)$. We first show that ℓ intersects C. Suppose not. Translate ℓ towards p until ℓ touches C at some point q. We claim that $q \in \text{int}(C)$. By this claim, ℓ is tangent to S_α at q. Observe that n_q lies on the plane through q and orthogonal to ℓ. This implies that $\angle(n_p, n_q) \geq \pi/2 - \angle(n_p, \ell) \geq \pi/4$, which contradicts Lemma 4.1.2(ii). We now prove the claim. Suppose that $q \in \partial C$. Then

q lies on the boundary of $B(p, kd)$ and between the medial balls at p. Refer to Figure 4.1. Let ℓ_p be the support line of n_p. Let ℓ' be the projection of ℓ onto the plane containing ℓ_p and q. Starting with Lemma 4.1.2(i), we have $\angle(pq, T_p) \leq \arcsin(kd) \leq 1/10$ as $kd \leq 1/20$. Also, $\angle(\ell_p, \ell') \leq \angle(\ell_p, \ell) \leq \pi/4$. Thus

$$\begin{aligned} \angle(pq, \ell') &= \pi/2 - \angle(\ell_p, \ell') - \angle(pq, T_p) \\ &\geq \pi/4 - 1/10. \end{aligned}$$

But then $d(p, \ell) \geq d(p, \ell') = kd \sin \angle(pq, \ell') > d$ as $k \geq 2$ and $\sin(\pi/4 - 1/10) > 1/2$, a contradiction.

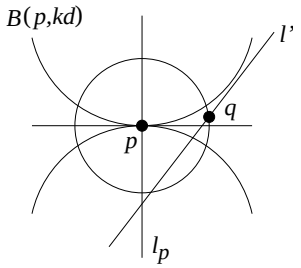

Figure 4.1: Illustration for the proof of Lemma 4.4.2.

Next, we show that ℓ intersects C at exactly one point. Suppose that there are two points $x, y \in \ell \cap C$. Note that $d(x, y) \leq 2kd$. By Lemma 4.1.2(i), $\angle(xy, T_x) \leq \arcsin(2kd) < 1/5$ as $kd \leq 1/20$. So $\angle(n_x, \ell) \geq \pi/2 - 1/5$, which implies that $\angle(n_x, n_p) \geq \pi/4 - 1/5 > 3kd$ as $kd \leq 1/20$. This contradicts Lemma 4.1.2(ii). □

Lemma 4.4.3 *Assume that $\delta \leq 1/2$. Let $p \in S_\alpha$ be a point. Let $d \leq 1/20$. Let Y be the infinite cylinder with radius d and axis aligned with n_p. Let C be the connected component of $Y \cap S_\alpha$ containing p. Let H be a plane passing p and n_p. Then $C \cap H$ is a curve segment, and C is a topological disk.*

Proof. First, we show that $C \subset S_\alpha \cap B(p, 2d)$. Consider any point q on C. Observe that the two balls of radius $1 - \delta$ tangent to S_α at p do not intersect $S_\alpha \backslash \{p\}$. As d is less than the radius $1 - \delta$ of the two balls, it is clear that $\angle(pq, T_p) < \pi/4$, which implies $d(p, q) < 2d$.

By Lemma 4.4.2, ℓ intersects $S_\alpha \cap B(p, 2d)$ at exactly one point. Since $C \subset S_\alpha \cap B(p, 2d)$, ℓ intersects C with at most one point. Suppose ℓ does not intersect C. Then we can translate ℓ toward n_p. At some point, ℓ will be tangent to some point z on C, which means $n_z = n_p$. However, we can prove that n_z and n_p make an acute angle, which then leads to a contradiction. As $d(p, z) \leq 2d$ and by Lemma 4.1.2(ii), $\angle(n_p, n_z) \leq 3d(p, z) \leq 6d < \pi/2$. So $\ell \cap C$ is exactly one point. Therefore $C \cap H$ is a curve segment, and C is a topological disk. □

Lemma 4.4.4 *Let p be a point on S_α. Let H be a plane passing through p making an acute angle with n_p. Let curve $F_\alpha = S_\alpha \cap H$. Then the radius of curvature of F_α at p is at least $(f(\tilde{p}) - \delta)\cos\angle(n_p, H)$.*

Proof. By Lemma 4.1.1, there are two tangent balls B_1, B_2 of S_α at p of radius $f(\tilde{p}) - \delta$ that do not intersect $S_\alpha \backslash \{p\}$. We show that there are two tangent disks D_1, D_2 of F_α at p of radius $(f(\tilde{p}) - \delta)\cos\angle(n_p, H)$ that do not intersect $F_\alpha \backslash \{p\}$.

Consider the plane H' containing n_p and the normal to H. Consider the cross-section of everything on H' in Figure 4.2. Let $D_1 = B_1 \cap H$. As B_1 does not

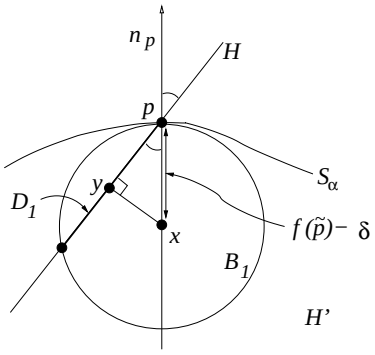

Figure 4.2: Illustration for the proof of Lemma 4.4.4.

intersect $S_\alpha \backslash \{p\}$, D_1 does not intersect $F_\alpha \backslash \{p\}$. Let x and y be the centers of B_1 and D_1, respectively. As $xy \perp py$, $\text{radius}(D_1) \geq d(p, y) = (f(\tilde{p}) - \delta)\cos\angle xpy = (f(\tilde{p}) - \delta)\cos\angle(n_p, H)$. □

Now we can do the proof for Lemma 4.3.3.

Proof of Lemma 4.3.3

First we establish the lower bound. Let H be a plane that contains the normal to S_α at p. Consider any point $q \in S_\alpha \cap Y \cap H$. Let C be the connected component of $S_\alpha \cap Y$ containing p. Let D be the projection area of C onto T_p. As C is a topological disk by Lemma 4.4.3, D is a disk. Thus area$(C) \geq$ area$(D) \geq \pi r^2$.

Then we will establish the upper bound. The radius of curvature at any point q on C is at least $1 - \delta$. Consider the projection of C onto T_p. Let J be the curve segment $H \cap C$ according to Lemma 4.4.3. Refer to Figure 4.3: $H \cap T_p$ is the x-axis,

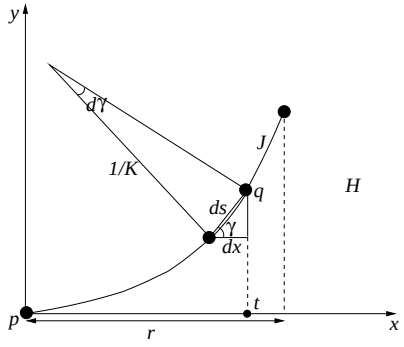

Figure 4.3: Illustration for the proof of Lemma 4.3.3.

and n_p is the y-axis. Let q be any point on J, and let t be its projection onto T_p. By Lemma 4.1.1, there are two tangent disks of J at p on H of radius $R = 1 - \alpha$ at p. As $r \leq R$, it is clear that $\angle(pq, T_p) \leq \pi/4$, which implies $d(q,t) \leq d(p,t) \leq r$. Thus $d(p,q) \leq \sqrt{2}r < 3r/2$. Then by Lemma 4.1.3, $d(\tilde{p}, \tilde{q}) \leq 2d(p,q) \leq 3r$. By Lemma 4.1.2(ii), $\angle(n_p, n_q) \leq 9r \leq 9/20 < \pi/6$, which implies $\angle(n_q, H) < \pi/6$ as n_p lies on H. By Lemma 4.4.4, the radius of curvature of J at q is at least $(1-\delta)\cos\angle(n_q, H) > (1-\delta)\cos\pi/6 = \frac{\sqrt{3}(1-\delta)}{2}$, which means the curvature $k(q)$ of any point on J is at most $\frac{2}{\sqrt{3}(1-\delta)}$. Let γ be the angle between the tangent of J at q and the horizontal. Let s be the arc length along J from p to q. Then as $ds = (\frac{1}{k(q)})d\gamma$,

$$d\gamma = k(q)ds = k(q)\sec\gamma dx \leq \frac{2}{\sqrt{3}(1-\delta)}\sec\gamma dx.$$

So $\frac{\sqrt{3}(1-\delta)}{2}\cos\gamma d\gamma = dx$. Thus by doing integration on both sides, $\frac{\sqrt{3}(1-\delta)}{2}\sin\gamma \leq x$,

which implies $\sin\gamma \leq \frac{2x}{\sqrt{3}(1-\delta)} \leq \frac{2r}{\sqrt{3}(1-\delta)}$ and $\cos\gamma \geq \sqrt{\frac{2}{3}}$ as $r \leq \frac{1-\delta}{20}$. Therefore, $ds = \sec\gamma dx \leq \sqrt{\frac{3}{2}}dx$.

Suppose H' is the plane obtained by rotating H through an angle $d\theta$ around n_p (i.e. the y-axis in Figure 4.3 and Figure 4.4). Let J' be the curve segment $H' \cap C$. Suppose Y' be a smaller cylinder concentric to Y and passing through a point q on J. Let q' be the intersection point $J' \cap Y'$. Refer to Figure 4.4. The arc length dt from q to q' along the boundary of $Y' \cap C$ is approximately $d(q, q')$.

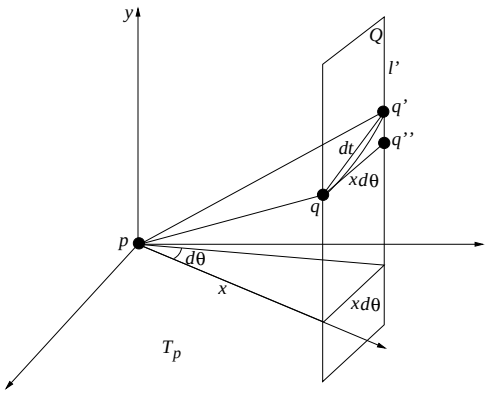

Figure 4.4: Illustration for the proof of Lemma 4.3.3.

Consider another plane Q passing through q and q' and perpendicular to T_p. Let ℓ' be the line through q' perpendicular to T_p. Let q'' be the projection of q onto ℓ'. Now $dt \approx d(q, q') = d(q, q'')\sec\angle q'qq'' = xd\theta\sec\angle q'qq''$. By Lemma 4.1.2 (i) and since dt is very small,

$$\begin{aligned}
\angle q'qq'' &= \angle(qq', T_p) \\
&\leq \angle(qq', T_q) + \angle(T_q, T_p) \\
&\leq \arcsin(d(q, q')) + \angle(n_p, n_q) \\
&\leq 2dt + \angle(n_p, n_q) \\
&< \pi/6.
\end{aligned}$$

Thus $dt \leq xd\theta\sec\frac{\pi}{6} = \frac{2}{\sqrt{3}}xd\theta$.

Refer to Figure 4.5. The surface area of C is $\int\int dsdt \leq \int\int(\sqrt{\frac{3}{2}}dx)(\frac{2}{\sqrt{3}}xd\theta) \leq \int_0^r\int_0^{2\pi}\sqrt{2}xd\theta dx = \int_0^r 2\sqrt{2}\pi x dx = [\sqrt{2}\pi x^2]_0^r = \sqrt{2}\pi r^2$. □

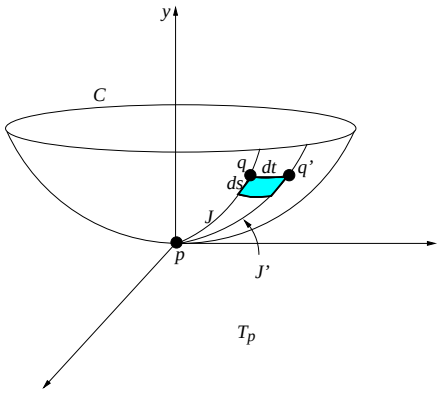

Figure 4.5: Illustration for the proof of Lemma 4.3.3.

CHAPTER 5

SURFACE RECONSTRUCTION

In this chapter, we will present an algorithm to reconstruct polygonal closed surfaces from noisy samples P obtained from a set of smooth closed surfaces. We assume P satisfies the conditions of the deterministic noise model as defined in the previous chapter. We show that the output polygonal reconstruction by our algorithm is faithful to the original surface S when $n = |P|$ is large enough.

Our surface reconstruction algorithm is described in section 5.1. Then we give an overview on the analysis of our faithfulness proof in Section 5.2, and we give the details of the analysis in Section 5.3, 5.4 and 5.5. We summarize in Section 5.6. Note that the proofs of some technical lemmas are given in the appendix.

5.1 Algorithm

Similar to the framework for the curve reconstruction algorithm in Chapter 3, our algorithm for surface consists of three main steps called Point estimation, Pruning and Surface reconstruction. Our algorithm first filters noise and then applies cocone to obtain a triangulated surface. The noise filtering procedure consists of two steps, Point estimation and Pruning. In the first step, which is the Point estimation step, new points are computed that are provably close to S. However, the distances of these new points from S can be much larger than the distances among them. Using all the new points to form a reconstruction would produce a highly rugged surface. In the second step, which is the Pruning step, some of the new points are pruned so that the interpoint distances in the remaining subset of new points are large relative to their distances from S. Now we can apply the cocone algorithm (or any faithful surface reconstruction algorithm for noiseless samples) to produce a provably good surface in the final Surface reconstruction step. We

remark that our algorithm does not estimate the model parameters $\epsilon, \epsilon_1, \epsilon_2$. It may be possible to do so, but we do not rely on any such estimate.

We further highlight the POINT ESTIMATION step. At each sample $p \in P$, we grow a ball neighborhood called *coarse neighborhood* until the noisy samples inside the neighborhood fit inside a sufficiently thin slab. This thin slab is an approximation of the tangent plane at $\tilde{p}$, i.e., the direction normal to the thin slab approximates n_p. However, due to noise, the size of the ball neighborhood is in the order of the noise amplitude δ. As the surface can bend quite a lot inside this ball neighborhood, the approximation error of the estimated normal is also in the order of δ. Thus we have a second phase to improve the estimate in a smaller neighborhood called *refined neighborhood* around p. We take some directions slightly different from the normal estimate due to the coarse neighborhood. We construct a narrow tube oriented in each direction. We trim each tube to the shortest cylinder containing all samples inside. Among all cylinders, the orientation of the shortest one provably approximates n_p well. We provide the algorithmic details below. Recall that $\omega > 0$, $\rho \geq 20$, and $0 < \gamma \leq 1/8$ are constants chosen in advance.

POINT ESTIMATION. For each noisy sample $p \in P$, we compute a point p^* that approximates $\tilde{p}$ as follows.

Initial ball. We compute a ball $\mathit{initial}(p)$ centered at p with $\ln^{\omega} n$ samples inside.

Coarse neighborhood. We initialize $\mathit{coarse}(p)$ to be the ball centered at p with radius $r_p = \text{radius}(\mathit{initial}(p))^{\gamma}$. We compute an infinite slab $\mathit{candidate}(p)$ that contains all samples inside $\mathit{coarse}(p)$. A slab of minimum width would be nice, but a 2-approximation suffices which can be computed in linear time [33]. We repeatedly double the size of $\mathit{coarse}(p)$ and maintain $\mathit{candidate}(p)$ until the radius of $\mathit{coarse}(p)$ is not less than $\rho \cdot \text{width}(\mathit{candidate}(p))$.

Refined neighborhood. Let $\mathcal{D}$ be the unit sphere of directions centered at p. We cover $\mathcal{D}$ by a set of caps as follows. Grid the boundary of

the bounding box of $\mathcal{D}$ into square cells of width $\sqrt{r_p}$. Project the grid vertices onto $\mathcal{D}$ towards the center of $\mathcal{D}$ and put a cap with angular radius $\sqrt{r_p}$ centered at each projected grid vertex. Let h_p be the line through p orthogonal to *candidate*(p). We examine all caps on $\mathcal{D}$. Let ℓ be the line through p and the center of a cap. If $\angle(h_p, \ell) > \pi/10$, examine the next cap. Otherwise, let $L_p(\ell)$ denote the infinite tube with axis ℓ and radius $r_p/4$. We trim $L_p(\ell)$ to obtain the shortest cylinder $C_p(\ell)$ that contains all noisy samples inside $L_p(\ell) \cap$ *coarse*(p). After going through all caps, we pick the shortest cylinder $C_p(\ell_p^*)$ generated. The line ℓ_p^* approximates n_p and the center p^* of $C_p(\ell_p^*)$ approximates $\tilde{p}$.

PRUNING. We scan the cylinder centers in an arbitrary order and select a subset C^*: when we select the center p^*, we delete all centers q^* not yet selected that satisfy $d(p^*, q^*) \leq \sqrt[8]{r_p/4}$.

SURFACE CONSTRUCTION. We run the cocone algorithm [5] on the centers in C^* and return the output surface, say N. One modification is that instead of using polar normals, we use the estimated normal ℓ_p^* for each $p^* \in C^*$. The details are provided in section 5.5.5.

We analyze the running time of the algorithm. PRUNING clearly takes $O(n^2)$ time. In SURFACE CONSTRUCTION, the most time-consuming step is the construction of the 3D Voronoi diagram which takes $O(n^2)$ time. INITIAL BALL takes $O(n)$ time for each sample. It remains to show that COARSE NEIGHBORHOOD and REFINED NEIGHBORHOOD take $O(n^{1+\gamma})$ time for each sample. We need a technical lemma on the relations among $n, \epsilon, \epsilon_1, \epsilon_2$, whose proof is given in the appendix. Recall that we are working with the deterministic noise model, so these relations are derived based on it.

Lemma 5.1.1 *$\epsilon, \epsilon_1 = O((\frac{\ln^\omega n}{n})^{2/9})$ and $\epsilon_2 = \Omega((\frac{1}{n})^{3/4})$.*

Lemma 5.1.2 *The algorithm runs in $O(n^{2+\gamma})$ time.*

Proof. In growing *coarse*(p), we spend $O(n)$ time to compute *candidate*(p) after each doubling. So the total time needed is $O(n \log(R_p/r_p))$, where R_p is the radius of the final *coarse*(p). By the model, radius(*initial*(p)) $\geq \epsilon_2$ which is $\Omega(1/n^{3/4})$ by Lemma 5.1.1. Thus, r_p = radius(*initial*(p))$^\gamma$ $= \Omega(1/n^{3\gamma/4})$. We will show in Lemma 5.3.2 that $R_p \leq 20\rho\delta + 2r_p$, which is $O(1 + r_p)$. Thus $\log(R_p/r_p) = O(\log n)$, implying that COARSE NEIGHBORHOOD takes $O(n \log n)$ time per sample. Thus it takes $O(n^2 \log n)$ total time.

In REFINED NEIGHBORHOOD, for each $L_p(\ell)$, collecting the samples inside $L_p(\ell) \cap$ *coarse*(p) and computing $C_p(\ell)$ takes $O(n)$ time. Let K be the set of caps on $\mathcal{D}$ considered in the algorithm for sample point p. Since the angular radius of each cap is $\sqrt{r_p}$, the area of each cap is $O(r_p)$. Thus the number of caps $|K|$ considered for sample point p is $O(1/r_p)$. So the total time to compute all $C_p(\ell)$ for the corresponding directions in K is $O(n/r_p)$. As $r_p = \Omega(1/n^{3\gamma/4})$, we have $n/r_p = O(n^{1+3\gamma/4}) = O(n^{1+\gamma})$. Thus REFINED NEIGHBORHOOD takes $O(n^{1+\gamma})$ time per sample. Thus it takes $O(n^{2+\gamma})$ total time.

On the other hand, as PRUNING and SURFACE RECONSTRUCTION takes $O(n^2)$ time. Hence the total running time is $O(n^{2+\gamma})$. □

5.2 Overview of analysis

In the consequent sections, we will prove the theoretical guarantees ensured by our algorithm, that are summarized in the following theorem.

Theorem 5.2.1 *Assume $\rho > 1$ and $\delta \leq 1/(1600\rho^2)$. Let $0 < \gamma \leq 1/8$ be a constant parameter. Given a noisy sampling P of S that satisfies our deterministic noise model, our algorithm constructs a a triangulated surface N in $O(n^{2+\gamma})$ time such that as n increases,*

(i) each vertex p^* of N converges to $\tilde{p}$,

(ii) the normal of each triangle T in N converges to the normal at $\tilde{p}$ for some vertex p^* of T, and

(iii) N is homeomorphic to S.

Not surprisingly, the framework of our proofs for faithfulness for the reconstructed surface is similar to the curve case in Chapter 3. We need to go through the analysis for the coarse neighborhood, the refined neighborhood, and the homeomorphism proof.

Consider the final coarse neighborhood $coarse(p)$ at a sample point p. Due to the criterion to sustain the growing of $coarse(p)$, $coarse(s)$ must have radius $\Theta(\rho\delta + r_p)$, where r_p is the radius of initial coarse ball at p. Next, we would like to argue that the normal of $candidate(s)$ approximates the surface normal at $\tilde{s}$. We claim that the deviation is actually $O(\rho\delta + r_p)$. We prove this by contradiction and assume that $candidate(s)$ is tilted a lot. Then a significant volume of $\mathbf{S} \cap coarse(p)$ lies outside $candidate(s)$. Our goal is to show that this particular volume contains a noisy sample so that a contradiction is established. The details are in Section 5.3.

As the normal estimation in the coarse neighborhood depends on the noise amplitude, and thus does not converge to zero when n increases. So we need to proceed to have a better estimate in the refined neighborhood. In the refined neighborhood at p, we rotate a thin cylinder of width $\Theta(r_p)$ centered at p. In the algorithm, the both bases of a cylinder are set to tightly bound all the sample points inside the cylinder. Our target cylinder axis direction ℓ_p^* is a direction so that the cylinder height is minimum. For efficiency, in our algorithm, we only seek for an approximate minimum height. The proof intuition is that when ℓ_p^* is close to the surface normal, the corresponding cylinder height will be quite small; otherwise the corresponding cylinder height will be large. Thus we can eliminate the latter option, and we can prove that ℓ_p^* is close to the surface normal. From this, we can then prove the center p^* of this cylinder of small height is close to the original surface S. The details are in Section 5.4.

Finally, to prove that the normals of the output triangles approximate well the surface normals and the reconstructed surface is homeomorphic to the original surface S, we adapts the proof technique in the cocone paper [5] for noiseless sample points. The main problem we need to get through is that in their original proof they

assume that all the sample points must be on the original surface S, which is not true in our case. The details are in Section 5.5.

5.3 Coarse neighborhood

We first bound the size of $coarse(p)$ which allows us to bound the approximation error of the initial estimate of n_p.

5.3.1 Radii of $initial(p)$ and $coarse(p)$

The next two lemmas bound the radii of $initial(p)$ and $coarse(p)$.

Lemma 5.3.1 *Then* $\Omega(\epsilon^{3/2}) = \epsilon_2 \leq \text{radius}(initial(p)) \leq \epsilon_1 = O(\epsilon)$. *Also,* $r_p^2 \geq \epsilon$ *and* $r_p = O(\epsilon^\gamma)$ *for sufficiently large* n.

Proof. By our deterministic model, as $initial(p)$ contains $\ln^\omega n$ samples, the radius of $initial(p)$ is less than $\epsilon_1 = O(\epsilon)$ and at least $\epsilon_2 = \Omega(\epsilon^{3/2})$. Then $r_p = (\text{radius}(initial(p)))^\gamma = O(\epsilon^\gamma)$.

Since $r_p \geq \text{radius}(initial(p))^\gamma \geq \epsilon_2^\gamma$ and $\epsilon_2 = \Omega(1/n^{3/4})$ by Lemma 5.1.1, we have $r_p = \Omega(1/n^{3\gamma/4}) = \Omega(1/n^{3/32})$ as $\gamma \leq 1/8$. By Lemma 5.1.1, $\epsilon = O((\frac{\ln^\omega n}{n})^{2/9})$. Thus for sufficiently large n, $r_p^2 \geq \epsilon$. □

Lemma 5.3.2 *Assume* $\rho > 1$ *and* $\delta \leq 1/(1600\rho^2)$. *Then*

$$\max\{2\sqrt{\rho}\delta, r_p\} \leq \text{radius}(coarse(p)) \leq 10\rho\delta + 2r_p$$

for sufficiently large n.

Proof. Let $R_p = \text{radius}(coarse(p))$. We first upper bound R_p. Suppose to the contrary that $R_p > 10\rho\delta + 2r_p$. Since the final $coarse(p)$ is obtained by repeated doubling $B(p, r_p)$. The algorithm must come across a ball B such that $5\rho\delta + r_p \leq \text{radius}(B) \leq 10\rho\delta + 2r_p$. Before reaching the final $coarse(p)$, we show that the

algorithm should have stopped and reported B as the final coarse ball, which would lead to contradiction.

Let x and y be points on S_δ^+ and S_δ^-, respectively, such that $\tilde{x} = \tilde{y} = \tilde{p}$. As $r_p = O(\epsilon^\gamma)$ by Lemma 5.3.1, $r_p \to 0$ as $n \to \infty$. As $\delta < 1/(1600\rho^2)$ and for sufficiently large n,

$$\text{radius}(B) \leq 10\rho\delta + 2r_p < (f(\tilde{p}) - \delta)/2.$$

Thus for any points $a, b \in B \cap S_\delta^+$, $d(a,b) \leq 2\,\text{radius}(B) < f(\tilde{p}) - \delta$. Let d_x be the maximum distance between $B \cap S_\delta^+$ and T_x. As $r_p < 1/(320\rho)$ for large n by Lemma 5.3.1 and $40\rho \leq 1/\sqrt{\delta}$, Lemma 4.1.2(i) implies that

$$\begin{aligned}
d_x &\leq (2\,\text{radius}(B))^2 \\
&\leq (20\rho\delta + 4\,r_p)^2 \\
&\leq (\sqrt{\delta}/2 + 4\,r_p)^2 \\
&\leq \delta/4 + (4\sqrt{\delta} + 16\,r_p)r_p \\
&\leq \delta/4 + (\frac{8}{40\rho} + \frac{1}{20\rho})r_p \\
&\leq \delta/4 + \frac{r_p}{4\rho}.
\end{aligned}$$

Similarly, the maximum distance d_y between $B \cap S_\delta^-$ and T_y is at most $\delta + \frac{r_p}{4\rho}$. Therefore, if we enclose the noisy samples in $B \cap \mathbf{S}$ using a slab L parallel to T_p, then $\text{width}(L) \leq d_x + d_y + 2\delta \leq 4\delta + \frac{r_p}{2\rho} \leq \text{radius}(B)/(2\rho)$. Since we compute a 2-approximation of the thinnest slab, the slab width computed by the algorithm for B is at most $\text{radius}(()B)/\rho$. But then, the algorithm should have stopped and reported B. This is a contradiction.

Next, we lower bound R_p. Clearly, $R_p \geq r_p$ by construction. We show that $R_p \geq 2\sqrt{\rho}\delta$. Let $B = B(p, R_p/\sqrt{\rho})$ and let $X = B(\tilde{p}, \delta)$. Note that $p \in X$ and X is tangent to S_δ^+ and S_δ^-. Lemma 4.1.1 implies that X is sandwiched between S_δ^+ and S_δ^- as $f(\tilde{p}) - \delta > \delta$.

Assume to the contrary that $R_p < 2\sqrt{\rho}\delta$. Then $\text{radius}(B) < 2\delta$. Since $p \in X$, $B \cap X$ contains a ball with radius $\text{radius}(B)/4$. The width of $\mathit{candidate}(p)$ is at most

$R_p/\rho = \text{radius}(B)/\sqrt{\rho}$. Thus, $(B \cap X) - \mathit{candidate}(s)$ contains a ball Y such that Y is empty, Y is sandwiched between S_δ^+ and S_δ^-, and

$$\begin{aligned} \text{radius}(Y) &\geq (\frac{1}{8} - \frac{1}{8\sqrt{\rho}}) \cdot \text{radius}(B) \\ &\geq R_p/(16\sqrt{\rho}). \end{aligned}$$

Since $R_p \geq r_p$ by construction, $\text{radius}(Y) \geq r_p/(16\sqrt{\rho}) \geq r_p^2$ for sufficiently large n. By Lemma 5.3.1, $r_p^2 \geq \epsilon$. But then Y contains a noisy sample, a contradiction. Hence $R_p \geq \min\{2\sqrt{\rho}\delta, r_p\}$. □

5.3.2 Rough normal estimate

We remark that the following analysis of the approximation error assumes that $r_p < 1$. Since $r_p = \text{radius}(\mathit{initial}(p))^\gamma \leq O(\epsilon^\gamma)$ and $\epsilon = O((\frac{\ln^\omega n}{n})^{2/9})$, the smaller γ is, the larger n needs to be so that $r_p < 1$. This is a tradeoff between the running time and the number of samples needed to achieve a small approximation error. Of course, as γ is a constant with respect to n, γ does not affect the asymptotic behavior of the algorithm. Next, we show that $\mathit{candidate}(p)$ is a rough approximation of T_p in Lemma 5.3.4. To do this, we need one more technical lemma (Lemma 5.3.3), which proof is deferred to appendix.

Lemma 5.3.3 *Assume that $\delta \leq 1/2$. Let $p \in S_\alpha$ be a point. Let $d \leq 1/20$. Let H be a plane passing p and n_p. Then $S_\alpha \cap B(p,d) \cap H$ is a curve segment, and $S_\alpha \cap B(p,d)$ is a topological disk.*

Lemma 5.3.4 *Assume $\rho > 1$ and $\delta \leq 1/(1600\rho^2)$. For sufficiently large n, the angle between T_p and $\mathit{candidate}(p)$ is at most $20\rho\delta + 4r_p + 4/\rho$.*

Proof. Let $p \in S_\alpha$. If T_p and $\mathit{candidate}(p)$ are parallel, we are done. Suppose not. Let H be the plane spanned by n_p and the line perpendicular to $\mathit{candidate}(p)$. Figure 5.1 shows the cross-section on H. Let L denote the strip $\mathit{candidate}(p) \cap H$. Note that $\text{width}(L) = \text{width}(\mathit{candidate}(p))$. Let ℓ be the line on H through p parallel to L. By

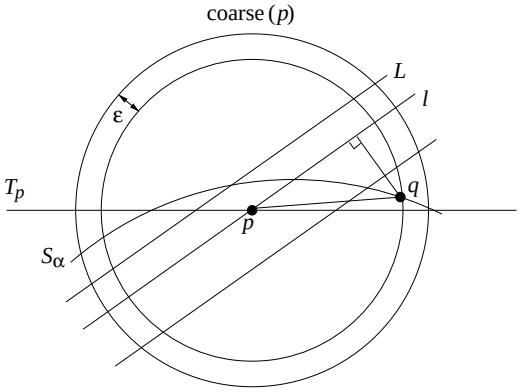

Figure 5.1: Illustration for the proof of Lemma 5.3.4.

Lemma 5.3.3, the boundary of $B(p, R_p - \epsilon)$ intersects the curve $S_\alpha \cap H \cap \mathit{coarse}\,(p)$ at exactly two points. Let q be one of them.

Since $B(q, \epsilon)$ contains some noisy sample, $B(q, \epsilon)$ must intersect $\mathit{candidate}\,(p)$ as $\mathit{candidate}\,(p)$ contains all noisy samples inside $\mathit{coarse}\,(p)$. Thus $B(q, \epsilon) \cap H$ must intersect L.

Assume to the contrary that $\angle(T_p, \ell) > 20\rho\delta + 4r_p + 4/\rho$. Since $R_p < f(\tilde{p}) - \delta$, Lemma 4.1.2(i) implies that $\angle(pq, T_p) \leq \arcsin(R_p - \epsilon) \leq \arcsin R_p$, which is at most $\arcsin(10\rho\delta + 2r_p)$ by Lemma 5.3.2. Then $\angle(pq, \ell) \geq \angle(T_p, \ell) - \arcsin(10\rho\delta + 2r_p) > 20\rho\delta + 4r_p + 4/\rho - 20\rho\delta - 4r_p = 4/\rho$. It follows that

$$d(q, \ell) \geq d(p, q) \cdot \sin\angle(pq, \ell) > (R_p - \epsilon) \cdot \frac{2}{\rho} = \frac{R_p}{\rho} + \epsilon + \frac{R_p - 2\epsilon - \rho\epsilon}{\rho}.$$

As $r_p^2 \geq \epsilon$ by Lemma 5.3.1 and $\sqrt{\epsilon} \leq 1/(\rho + 2)$ for large n by Lemma 5.1.1, we have $R_p \geq r_p \geq \sqrt{\epsilon} \geq (\rho+2)\epsilon$. This implies that $d(q, \ell) > R_p/\rho + \epsilon \geq \mathrm{width}(L) + \epsilon$, which contradicts the fact that $B(q, \epsilon) \cap H$ intersects L. □

5.4 Refined neighborhood and point convergence

By Lemma 5.3.4, since $10\rho\delta \leq \pi/12$ as $\delta \leq 1/(1600\rho^2)$, the angle between T_p and $\mathit{candidate}\,(p)$ is at most $\pi/12$ for sufficiently large n. Since our algorithm enforces

the angle between a tube axis ℓ and the direction orthogonal to *candidate*(p) is at most $\pi/10$, $\angle(n_p, \ell) \leq \pi/12 + \pi/10 < \pi/5$. Recall that the radii of the tubes $L_p(\ell)$ and cylinders $C_p(\ell)$ are set to $r_p/4$. We want to show that REFINED NEIGHBORHOOD gives a provably good estimate of n_p. We need the following several technical lemmas, whose proofs are given in the appendix.

Lemma 5.4.1 *Let ℓ be an arbitrary line through a sample point p such that $\angle(n_p, \ell) \leq \pi/5$. For any $\alpha \in [-\delta, \delta]$, $L_p(\ell) \cap coarse(p) \cap S_\alpha$ is a topological disk inside $B(x, r_p/2)$, where $x = \ell \cap coarse(p) \cap S_\alpha$.*

The following lemma bounds some distances from p to S_α in terms of $\angle(n_p, \ell)$.

Lemma 5.4.2 *Let ℓ be an arbitrary line through a sample point p such that $\theta = \angle(n_p, \ell) \leq \pi/5$. For any α, let $p_1 \in S_\alpha$ with $\tilde{p_1} = \tilde{p}$. Let $x = \ell \cap coarse(p) \cap S_\alpha$. Then (i) $d(p_1, x) \leq \theta d(p, p_1)/2$, (ii) $d(p, x) \leq \frac{d(p,p_1)+\theta^2 d(p,p_1)^2/4}{\cos\theta}$, and (iii) $d(p, x) \geq \frac{d(p,p_1)-\theta^2 d(p,p_1)^2/4}{\cos\theta}$.*

The following lemma is for proving the point convergence later.

Lemma 5.4.3 *Let ℓ be a line through a sample point p such that $\theta = \angle(n_p, \ell) \leq \pi/5$. Let e be the segment $\ell \cap \mathbf{S} \cap coarse(p)$. Let q be the point $e \cap S$. Then $d(q, m) \leq 2\delta^2\theta^2/\cos\theta$, where m is the midpoint of e.*

Proof. Suppose $e = q_1q_2$ where $q_1 \in S_\delta^+, q_2 \in S_\delta^-$. Let p_1p_2 denote the normal segment passing through p, where $p_1 \in S_\delta^+, p_2 \in S_\delta^-$. Let $t_1 = \ell \cap T_{p_1}$ and $t_2 = \ell \cap T_{p_2}$. And let $t = \ell \cap T_{\tilde{p}}$. Refer to Figure 5.2. Note that t is the midpoint of the segment t_1t_2. By Lemma 5.4.2 (ii) & (iii) and the fact that $d(p, t_1) = d(p, p_1)/\cos\theta$, we get

$$|d(p, q_1) - d(p, t_1)| \leq \frac{\theta^2 d(p, p_1)^2}{4\cos\theta},$$

which is at most $\frac{\theta^2\delta^2}{\cos\theta}$ as $d(p, p_1) \leq 2\delta$. Since p, q_1, t are collinear, we conclude that

$$d(t_1, q_1) \leq |d(p, q_1) - d(p, t_1)| \leq \frac{\theta^2\delta^2}{\cos\theta}.$$

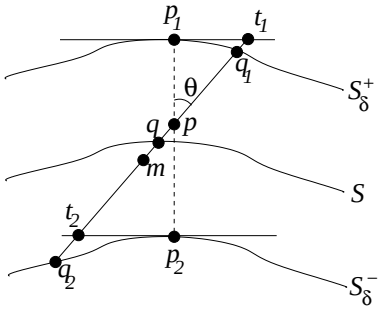

Figure 5.2: Illustration for the proof of Lemma 5.4.3.

Similarly we can prove $d(t_2, q_2) \leq \delta^2\theta^2/\cos\theta$. Thus

$$d(t, m) \leq \delta^2\theta^2/\cos\theta. \tag{5.1}$$

The same argument also shows that $|d(p, q) - d(p, t)| \leq \frac{\theta^2\delta^2}{\cos\theta}$. Since p, q, t are collinear, we conclude that

$$d(t, q) \leq |d(p, q) - d(p, t)| \leq \frac{\theta^2\delta^2}{\cos\theta}. \tag{5.2}$$

Hence, by 5.1 and 5.2, $d(q, m) \leq d(t, m) + d(t, q) \leq 2\delta^2\theta^2/\cos\theta$. □

5.4.1 Height bounds of cylinders

We prove upper and lower bounds on the heights of cylinders tested by the algorithm.

Lemma 5.4.4 *Let $\theta = \angle(n_p, \ell) \leq \pi/5$. Assume that $\theta \leq \sqrt{r_p}$. Let p be the point $\ell \cap S \cap coarse(p)$. Let p^* be the center of $C_p(\ell)$. Then the height of $C_p(\ell)$ is at most $\frac{2\delta+\theta^2\delta^2}{\cos\theta} + 2r_p\sqrt{r_p}$; and $d(p^*, q) \leq \frac{2\theta^2\delta^2}{\cos\theta} + r_p\sqrt{r_p}$.*

Proof. Let $e = q_1q_2$ be the segment $\ell \cap \mathbf{S} \cap coarse(p)$, where $q_1 \in S_\delta^+, q_2 \in S_\delta^-$. Let B_1, B_2 be the upper and lower bases of $C_p(\ell)$ respectively. Assume that B_i lies on the same side of q as q_i. Let $b_1 = \ell \cap B_1$ and $b_2 = \ell \cap B_2$. Let p_1p_2 be the normal segment passing through p where $p_1 \in S_\delta^+, p_2 \in S_\delta^-$. Note that T_p, T_{p_1}, T_{p_2} are parallel. Refer to Figure 5.3

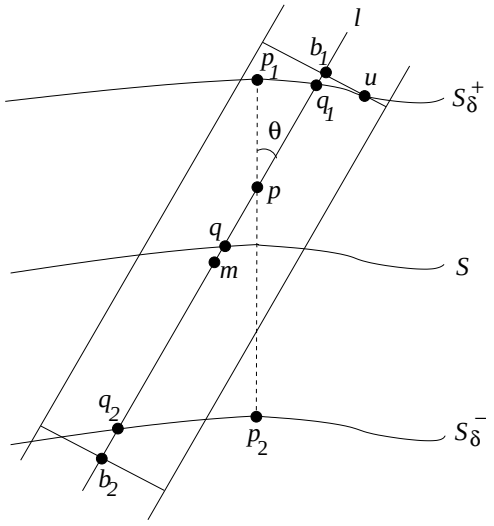

Figure 5.3: Illustration for the proof of Lemma 5.4.4.

We first show that $d(q_1, b_1) \leq r_p\sqrt{r_p}$. There are two cases to consider.

Case 1: b_1 lies between q_1 and q. Since $B(q_1, \epsilon)$ contains a sample point, we have $d(q_1, b_1) \leq \epsilon$, which is at most r_p^2 by Lemma 5.3.1.

Case 2: q_1 lies between b_1 and q. In this case, B_1 must intersect the patch $L_p(\ell) \cap S_\delta^+ \cap coarse(p)$, otherwise we could have moved B_1 close to q to decrease the height of $C_p(\ell)$. Let u be a point at the intersection. So $d(q_1, b_1) \leq r_p/4 \cdot \tan \angle q_1 u b_1$. We bound $\angle q_1 u b_1$ as follows.

$$\begin{aligned} \angle q_1 u b_1 &\leq \angle(q_1 u, T_{q_1}) + \angle(n_{q_1}, \ell) \\ &\leq \angle(q_1 u, T_{q_1}) + \angle(n_{p_1}, n_{q_1}) + \angle(n_{p_1}, \ell) \\ &= \angle(q_u, T_{q_1}) + \angle(n_{p_1}, n_{q_1}) + \theta. \end{aligned}$$

Since $d(q_1, u) \leq r_p/2$ by Lemma 5.4.1, Lemma 4.1.2 (ii) implies that $\angle(q_1 u, T_{q_1}) \leq \arcsin(r_p/2) \leq r_p$. By Lemma 4.1.2 (ii), $\angle(n_{p_1}, n_{q_1}) \leq 3d(p_1, q_1)$. By Lemma 5.4.2 (i), $d(p_1, q_1) \leq \theta d(p, p_1)/2 \leq \delta\theta$. So $\angle(n_{p_1}, n_{q_1}) \leq 3\delta\theta$. Therefor $\angle q_1 u b_1 \leq r_p + 3\delta\theta + \theta \leq 2\sqrt{r_p}$ as $\theta \leq \sqrt{r_p}$ by assumption. Hence $d(q_1, b_1) \leq r_p/4 \cdot \tan(2\sqrt{r_p}) \leq r_p\sqrt{r_p}$ for sufficiently small r_p.

Similarly we can prove $d(q_2, b_2) \leq r_p\sqrt{r_p}$. It follows that

$$\begin{aligned} d(p, b_1) &\leq d(p, q_1) + d(q_1, b_1) \\ &\overset{Lemma\ 5.4.2(ii)}{\leq} \frac{d(p, p_1) + \theta^2 d(p, p_1)^2/4}{\cos\theta} + r_p\sqrt{r_p}. \end{aligned}$$

Similarly, we also have

$$d(p, b_2) \leq \frac{d(p, p_2) + \theta^2 d(p, p_2)^2/4}{\cos\theta} + r_p\sqrt{r_p}.$$

Hence the height of $C_p(\ell)$ is $d(b_1, b_2) \leq \frac{2\delta+\theta^2\delta^2}{\cos\theta} + 2r_p\sqrt{r_p}$.

Next, we upper bound $d(p^*, q)$. Let m be the midpoint of e. As p^* is the midpoint of b_1b_2, we have $d(p^*, m) \leq r_p\sqrt{r_p}$. On the other hand, by Lemma 5.4.3 (i), $d(q, m) \leq 2\delta^2\theta^2/\cos\theta$. Thus by triangle inequality, we have

$$\begin{aligned} d(p^*, q) &\leq d(q, m) + d(p^*, m) \\ &\leq 2\delta^2\theta^2/\cos\theta + r_p\sqrt{r_p}. \end{aligned}$$

□

Lemma 5.4.5 *Let $\theta = \angle(n_p, \ell) \leq \pi/5$. Assume that $\theta > 32\sqrt{r_p}$. Then the height of $C_p(\ell)$ is at least $\frac{2\delta-\theta^2\delta^2}{\cos\theta} + 2r_p\sqrt{r_p}$.*

Proof. Assume that ℓ is vertical. Let H be the plane spanned by ℓ and n_p. Let p_1 be the point on S_δ such that $\tilde{p_1} = \tilde{p}$. Let ξ denote $S_\delta \cap L_p(\ell) \cap H$. By Lemma 5.3.3, ξ is a curve segment. Let a and b be the endpoints of ξ.

We first analyze the directions of the normals of ξ on H. Consider any point $z \in \xi$. Let n'_z be the projection of n_z onto H. Note that n'_z is the normal of ξ at z on H. Observe that $\angle(n_p, n'_z) \leq \angle(n_p, n_z) = \angle(n_{p_1}, n_2)$. By Lemmas 5.4.1 and 5.4.2(i), $d(p_1, z) \leq \theta\delta + r_p/2$. Lemma 4.1.2 (i) implies $\angle(n_{p_1}, n_z) \leq \arcsin(\theta\delta + r_p/2) \leq 2\theta\delta + r_p$. So $\angle(n_p, n'_z) \leq 2\theta\delta + r_p$.

Assume that $\theta > 32\sqrt{r_p}$. We argue that ξ strictly increases or decreases from a to b. If n_p points to the left of ℓ, then n'_z also points to the left of ℓ, making an angle $(1-2\delta)\theta - r_p > 0$ as $\theta > 32\sqrt{r_p}$. So ξ strictly increases from a to b. Conversely, if n_p points to the right of ℓ, then ξ strictly decreases from a to b.

W.l.o.g., assume that ξ strictly increases from a to b. Refer to Figure 5.4. Let q_1 be the point $\ell \cap S_\delta$. Let g be the horizontal line through q_1. Let h be the vertical

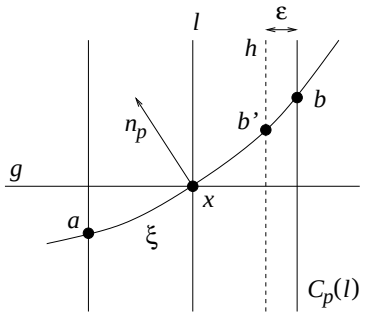

Figure 5.4: Illustration for the proof of Lemma 5.4.5.

line at distance ϵ from b. Let b' be the point $h \cap \xi$. Then

$$\begin{aligned} d(b', g) &\geq (r_p/4 - \epsilon)\tan((1-2\delta)\theta - r_p) \\ &\geq (r_p/4 - \epsilon)((1-2\delta)\theta - r_p). \end{aligned}$$

As $\epsilon \leq r_p^2$, $r_p/4 - \epsilon \geq r_p(1/4 - r_p)$ which is at least $r_p/8$ for sufficiently large n. Since $\theta > 32\sqrt{r_p}$, $(1-2\delta)\theta - r_p > 16\sqrt{r_p}$ when r_p is small. We conclude that $d(b', g) \geq 2r_p\sqrt{r_p}$.

Since the ball $B(b', \epsilon)$ contains a noisy sample, the upper end of $C_p(\ell)$ must be at distance at least $d(b', g) - \epsilon$ from g, which is at least $2r_p\sqrt{r_p} - \epsilon \geq r_p\sqrt{r_p}$ as $\epsilon \leq r_p^2$. By Lemma 5.4.2(iii), the distance from p to the upper end of $C_p(\ell)$ is at least $\frac{d(p,p_1) - \theta^2 d(p,p_1)^2/4}{\cos\theta} + r_p\sqrt{r_p}$. We can derive a similar lower bound for the distance from p to the lower end of $C_p(\ell)$. Hence the height of $C_p(\ell)$ is at least $\frac{2\delta - \theta^2\delta^2}{\cos\theta} + 2r_p\sqrt{r_p}$. □

5.4.2 Point convergence

We are now ready to show the refined cylinder direction ℓ_p^* and the cylinder center p^* approximate normal n_p and surface point $\tilde{p}$ well.

Lemma 5.4.6 (Point convergence) *Assume that $\delta \leq 1/(1600\rho^2)$. Let p^* and ℓ_p^* be the center point and approximate normal direction computed for p. Then $\angle(n_p, \ell_p^*) = O(\sqrt{r_p})$ and $d(\tilde{p}, p^*) = O(\sqrt{r_p})$.*

Proof. Consider the cylinder $C_p(\ell_1)$ computed by the algorithm. Let h be its height and let $\theta = \angle(n_p, \ell_p^*)$. Consider the caps covering the sphere of directions centered at p in REFINED NEIGHBORHOOD. n_p projects to a point inside some cap. Let ℓ' be the line through the center of this cap and p. So $C_p(\ell')$ is the corresponding cylinder. Let h' be its height and let $\theta' = \angle(n_p, \ell')$. Note that $\theta' \leq \sqrt{r_p}$.

Assume to the contrary that $\theta > 32\sqrt{r_p}$. By Lemmas 5.4.4 and 5.4.5, $h - h' \geq C - C'$, where

$$C = \frac{2\delta}{\cos\theta\cos\theta'}(\cos\theta' - \cos\theta), \text{ and}$$

$$C' = \frac{\delta^2}{\cos\theta\cos\theta'}(\theta^2\cos\theta' + \theta'^2\cos\theta) \leq \frac{2\delta^2\theta^2}{\cos\theta\cos\theta'}.$$

Since $\theta \geq 32\theta'$,

$$\begin{aligned}
\cos\theta' - \cos\theta &= 2\sin\frac{\theta' + \theta}{2}\sin\frac{\theta - \theta'}{2} \\
&\geq 2\sin\frac{\theta}{2}\sin\frac{\theta}{4} \\
&\geq 2\cdot\frac{\theta}{4}\cdot\frac{\theta}{8} \\
&\geq \theta^2/16.
\end{aligned}$$

Thus,

$$\begin{aligned}
C - C' &\geq \frac{\delta\theta^2}{8\cos\theta\cos\theta'} - \frac{2\delta^2\theta^2}{\cos\theta\cos\theta'} \\
&\geq \frac{\delta\theta^2}{8\cos\theta\cos\theta'}(1 - 16\delta) \\
&> 0
\end{aligned}$$

as $\delta \leq 1/(1600\rho^2)$. So $h - h' > 0$. This contradicts the fact that $C_p(\ell_p^*)$ has the minimum height. Hence $\theta \leq 32\sqrt{r_p}$.

Let q be the point $\ell_p^* \cap S$. By Lemma 5.4.2 (i), $d(\tilde{p}, q) \leq \theta d(p, \tilde{p})/2$, which is at most $\delta\theta$ as $d(p, \tilde{p}) \leq 2\delta$. By Lemma 5.4.4, $d(p^*, q) \leq \frac{2\theta^2\delta^2}{\cos\theta} + r_p\sqrt{r_p}$. Thus $d(\tilde{p}, p^*) \leq d(\tilde{p}, q) + d(p^*, q) = r_p\sqrt{r_p} + O(\theta) = O(\sqrt{r_p})$ as $\theta \leq 32\sqrt{r_p}$. □

5.5 Triangle normal and homeomorphism

Since the selected center points in C^* are not on the surface S, we cannot directly use the proofs in the cocone paper [5] for noiseless sample points. Instead we adapts their technique for our homeomorphism proof. We first need a concept called *Restricted Delaunay triangulation* on S, which was defined in say [23]. We will describe it first in the subsection below.

5.5.1 Restricted Delaunay triangulation

The *Voronoi diagram* V_{C^*} of C^* contains the *Voronoi cells* V_{p^*} for all points $p^* \in C^*$, where V_{p^*} is defined to be

$$\{x \in R^3 : d(x, p^*) \leq d(x, q^*) \text{ for any } q^* \in C^* \text{ and } q^* \neq p^*\}.$$

The *Delaunay triangulation* D_{C^*} of C^* has a edge p^*q^* if $V_{p^*} \cap V_{q^*}$ is a face, has a triangle $p^*q^*r^*$ if $V_{p^*} \cap V_{q^*} \cap V_{r^*}$ is an edge, and has tetrahedron $p^*q^*r^*s^*$ if $V_{p^*} \cap V_{q^*} \cap V_{r^*} \cap V_{s^*}$ is a point.

The restriction of V_{C^*} to the surface S is called the *restricted Voronoi diagram* of C^* on S, denoted as $V_{C^*}|_S$. It contains all the *restricted Voronoi cells* $V_{p^*}|_S = V_{p^*} \cap S$. The dual of these restricted Voronoi cells defines the *restricted Delaunay triangulation* $D_{C^*}|_S$. Specifically, $D_{C^*}|_S$ has an edge p^*q^* if and only if $V_{p^*}|_S \cap V_{q^*}|_S \neq \emptyset$, and a triangle $p^*q^*r^*$ if and only if $V_{p^*}|_S \cap V_{q^*}|_S \cap V_{r^*}|_S \neq \emptyset$. Assuming that S does not pass through a Voronoi vertex in general position, there is no tetrahedron in $D_{C^*}|_S$.

Edelsbrunner and Shah [23] showed that the underlying space of $D_{C^*}|_S$ is homeomorphic to S if the following *closed ball property* holds: each $V_{p^*}|_S$ is a topological disk, each nonempty pairwise intersection $V_{p^*}|_S \cap V_{q^*}|_S$ is a curve segment, and each nonempty triple intersection $V_{p^*}|_S \cap V_{q^*}|_S \cap V_{r^*}|_S$ is a single point.

5.5.2 Overview for homeomorphism

Now let's have an overview on how we proceed to prove the homeomorphism.

In the algorithm, all the center points are scanned to select a subset C^* of center points: when a center point p^* is selected, we delete all unselected center points u^* such that $d(p^*, u^*) \leq r_p^{1/8}$. We prove that the output manifold by running the cocone algorithm on C^* is a faithful reconstruction of S. We assume the points in C^* are in general position. In practice most Delaunay triangulation codes simulate general position, so this is not an unreasonable assumption.

In the noiseless case, Amenta, Choi, Dey, and Leekha [5] prove that the cocone algorithm returns a faithful reconstruction. The key step is proving that the normal of each triangle converges to the normal at some triangle vertex. Based on this result, one can then show that the restricted Delaunay triangles are not removed during the manifold extraction. This is important as it guarantees that there is some underlying manifold to be extracted. Furthermore, the convergence of triangle normals implies that the output surface is locally flat, which is instrumental to proving the homeomorphism.

The main tool in proving the homeomorphism and the convergence of normals of triangles is that there are two large tangent balls on opposite sides of S that touch S at each noiseless sample. The radii of these tangent balls are at least the local feature size of the sample, and these tangent balls do not contain other noiseless samples. We claim that an analogous result holds in the noisy case. So the proofs in [5] can be adapted here.

Specifically, let p^* be the center point computed for a noisy sample, and let $p^* \in S_\alpha$. We claim that there are two large tangent balls B_1 and B_2 on opposite sides of S_α that touch S_α at p^*. The radii of B_1 and B_2 are at least c for some constant $0 < c < 1$. Note that c is very large compared to the inter-center distances which is $O(r_p^{1/8})$ after pruning. We do not need the radii of B_1 and B_2 to be proportional to the local feature size because we use uniform sampling.

By the point convergence lemma, the selected center points C^* lie inside a thin shell. We will look at this and some basic properties of the selected centers points C^* in Section 5.5.3. Then in Section 5.5.4, we will prove that there are two large empty tangent balls at each center point p^*. With this lemma, we prove that the union of all restricted Delaunay triangles is homeomorphic to S in Section 5.5.6.

In Section 5.5.5, we review the cocone reconstruction algorithm. Finally in Sections 5.5.7, 5.5.8 and 5.5.9, we prove the properties on the sizes and normals of the reconstructed triangles and the homeomorphism between the reconstruction and the original surface S.

5.5.3 Thin shell

The point convergence lemma (Lemma 5.4.6) immediately implies the following lemma, which says that the center points are close to the surface S, and the estimated normals are close to the surface normals.

Lemma 5.5.1 *For any center point p^*, $d(\tilde{p^*}, \tilde{p}) = O(\epsilon^{\gamma/2})$ and $\angle(n_{p^*}, n_p) = O(\epsilon^{\gamma/2})$. Moreover, $d(p^*, \tilde{p^*}) = O(\epsilon^{\gamma/2})$.*

Proof. By Lemma 5.4.6 and Lemma 5.3.1, $d(p^*, \tilde{p}) = O(\sqrt{r_p}) = O(\epsilon^{\gamma/2})$. As $\tilde{p^*}$ is the closest point on S to p^*, $d(p^*, \tilde{p^*}) \leq d(p^*, \tilde{p}) = O(\epsilon^{\gamma/2})$. Thus $d(\tilde{p^*}, \tilde{p}) \leq d(p^*, \tilde{p^*}) + d(p^*, \tilde{p}) = O(\epsilon^{\gamma/2})$. By Lemma 4.1.2 (ii), $\angle(n_{p^*}, n_p) \leq 3d(\tilde{p^*}, \tilde{p}) = O(\epsilon^{\gamma/2})$.

Moreover, $d(p^*, \tilde{p^*}) \leq d(p^*, \tilde{p}) = O(\epsilon^{\gamma/2})$. ∎

By this lemma, we have $d(p^*, \tilde{p^*}) = O(\epsilon^{\gamma/2})$ for any $p^* \in C^*$. Let $\Delta = \max_{\{p^* \in C^*\}} d(p^*, \tilde{p^*})$, which is $O(\epsilon^{\gamma/2})$. Let $\mathbf{S}_\Delta$ be the volume between S_Δ^+ and S_Δ^-. Then all center points in C^* lie inside $\mathbf{S}_\Delta$. The following lemma says that the center points C^* in $\mathbf{S}_\Delta$ are dense.

Lemma 5.5.2 *For each point $q \in \mathbf{S}_\Delta$, there is a center point $p^* \in C^*$ such that $d(p^*, q) = O(\epsilon^{\gamma/8})$.*

Proof. The ball $B(q, \epsilon)$ contains a sample point $s \in P$ by our deterministic noise model. Then $d(q, s^*) \leq d(q, s) + d(s, \tilde{s}) + d(\tilde{s}, s^*) \leq \epsilon + (\Delta + \epsilon) + d(\tilde{s}, s^*) \leq O(\epsilon^{\gamma/2}) + d(\tilde{s}, s^*)$. Since $d(\tilde{s}, s^*) \leq d(\tilde{s}, \tilde{s^*}) + d(\tilde{s^*}, s^*) = O(\epsilon^{\gamma/2})$ by Lemma 5.5.1, we have $d(q, s^*) = O(\epsilon^{\gamma/2})$.

If $s^* \in C^*$, then we are done. Otherwise s^* is decimated when a center point p^* is selected. This means $d(s^*, p^*) \leq \sqrt[8]{r_p} = O(\epsilon^{\gamma/8})$. Thus $d(q, p^*) \leq d(q, s^*) + d(s^*, p^*) = O(\epsilon^{\gamma/8})$. □

Corollary 5.5.1 *For any $p^* \in C^*$, $V_{p^*}|_{S_alpha}$ is contained in a ball of radius $O(\epsilon^{\gamma/8})$ centered at p^*, where $\alpha \leq \Delta$.*

Next, we prove that after pruning, the distance between the selected center points is large when compared with Δ.

Lemma 5.5.3 *Let $p^*, q^* \in C^*$. Then $d(p^*, q^*) \geq \epsilon_2^{\gamma/8}$.*

Proof. Assume that, in the algorithm, p^* is selected before q^*. So $d(p^*, q^*) > r_p^{1/8}$. Recall that $r_p = \text{radius}(\mathit{initial}(p))^\gamma$ and $\text{radius}(\mathit{initial}(p)) \geq \epsilon_2$ by Lemma 5.3.1. So $d(p^*, q^*) \geq \epsilon_2^{\gamma/8}$. □

5.5.4 Twin empty tangent balls

We then prove that for any center point $p^* \in S_\alpha$, where $\alpha \leq \Delta$, there are twin balls tangent to S_α at p^* and empty of the center points in C^*. This lemma allows us to adapt many proofs from the cocone paper [5].

Lemma 5.5.4 *Suppose $p^* \in C^*$ is on S_α. For sufficiently large n, there are two balls tangent to S_α at p^* of radius $1/10$ on opposite sides of S_α such that their interiors do not contain any center point in C^*.*

Proof. We will consider only one side of S_α as the arguments for the other side are symmetric. Let m be a point at unit distance from $\tilde{p^*}$ in the normal direction. Let $M = B(m, 1-\Delta)$. Since $f_{\min} = 1$, the following holds: M is tangent to $\mathbf{S}_\Delta$, M is free from intersecting the interior of $\mathbf{S}_\Delta$, and M is empty of center points in C^*. Let q be the tangent point between M and S_Δ. We consider another ball $B = B(p^*, \epsilon_2^{\gamma/8})$, which is empty of center points in C^* by Lemma 5.5.3.

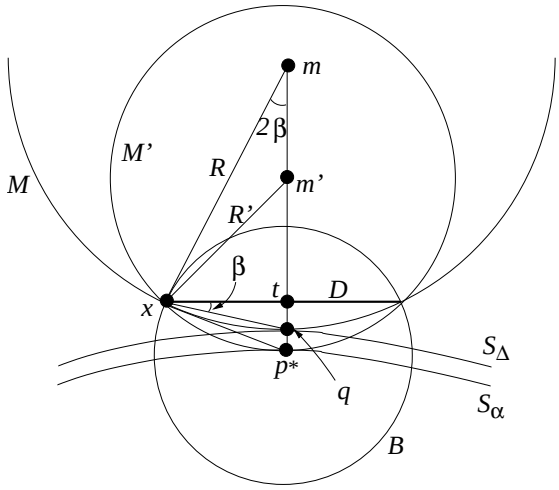

Figure 5.5: Illustration for the proof of Lemma 5.5.4.

Refer to Figure 5.5. Let D be the circle $\partial B \cap \partial M$. Let t be the center of D. Note that t lies on p^*m. We push M through circle D towards p^* until it touches p^* to produce a smaller ball $M' = B(m', r')$. We show that M' satisfies the lemma requirement by showing that $r' \geq 1/10$. Note that $m' \in p^*m$. So $M' \subset B \cup M$, which implies that M' does not contain any center point in C^*.

Consider a point x on the circle D. Note that $d(x, m') = r'$. As $\epsilon = O(\epsilon_2^{2/3})$, we have

$$d(x, q) \leq d(p^*, x) + d(p^*, q) \leq \epsilon_2^{\gamma/8} + 2\Delta.$$

Since $\Delta = O(\epsilon^{\gamma/2})$ and $\epsilon_2 = \Omega(\epsilon^{2/3})$, we have $2\Delta = O((\epsilon_2^{3/2})^{\gamma/2}) = O(\epsilon_2^{3\gamma/4})$, which is less than $\epsilon_2^{\gamma/8}$ for sufficiently large n. So $d(x, q) \leq 2\epsilon_2^{\gamma/8}$.

Let $\beta = \angle qxt$. Note that $\angle p^*mx = 2\beta$. Then $\sin\beta = \frac{d(x,q)/2}{1-\Delta} \leq \frac{\epsilon_2^{\gamma/8}}{1/2} = 2\epsilon_2^{\gamma/8}$. Thus

$$d(q, t) = d(x, q)\sin\beta \leq (2\epsilon_2^{\gamma/8})^2 = 4\epsilon_2^{\gamma/4}.$$

It follows that

$$\sin\angle p^*xt = \frac{d(p^*, q) + d(q, t)}{d(p^*, x)} \leq \frac{2\Delta + 4\epsilon_2^{\gamma/4}}{\epsilon_2^{\gamma/8}} = \frac{O(\epsilon_2^{3\gamma/4}) + 4\epsilon_2^{\gamma/4}}{\epsilon_2^{\gamma/8}}.$$

As $O(\epsilon_2^{3\gamma/4}) \leq \epsilon_2^{\gamma/4}$ for sufficiently large n,

$$\sin \angle p^* x t \leq \frac{5\epsilon_2^{\gamma/4}}{\epsilon_2^{\gamma/8}} = 5\epsilon_2^{\gamma/8}.$$

Hence $r' = \frac{d(p^*,x)/2}{\sin \angle p^* x t} \geq \frac{\epsilon_2^{\gamma/8}/2}{5\epsilon_2^{\gamma/8}} \geq 1/10.$ □

Consider a vertex $p^* \in S_\alpha$ of a restricted Delaunay triangle T. Since there are twin big balls tangent to S_α at p^* by the previous lemma, and the circumradius of a restricted Delaunay triangle is $O(\epsilon^{\gamma/8})$ by Corollary 5.5.1, we can prove the following lemma by using almost the same argument in Lemma 7 in [2]. The details of its proof is in the appendix.

Lemma 5.5.5 *Let τ be a restricted Delaunay triangle, and n_τ be the line normal to τ. Then $\angle(n_\tau, n_{p^*}) = O(\epsilon^{\gamma/8})$ for any vertex p^* of τ.*

With this lemma, by using almost the same arguments as Theorem 2 in [2], we have the following lemma. The details of its proof is in the appendix.

Lemma 5.5.6 *The union of restricted Delaunay triangles is homeomorphic to S.*

5.5.5 Review of the cocone algorithm

We will first review the cocone algorithm [5] because some understanding is needed to prove the homeomorphism.

First of all, the Delaunay triangulation of C^* is computed. For any point $p^* \in C^*$, let $e = w_1 w_2$ be an edge in V_{p^*}. Suppose ℓ_p^* be the estimated normal by previous algorithm at p^*. Let $\theta = \pi/8$. We then compute $\angle(p^* w_1, \ell_p^*)$ and $\angle(p^* w_2, \ell_p^*)$, and check if the range of angles between these two angles intersect $[\pi/2 - \theta, \pi/2 + \theta]$. If it does, it actually means the Voronoi edge e intersects the complement of the double cone with apex p^* and making an angle of $\pi/2 - \theta$ with axis ℓ_p^*, and we mark e. A Delaunay triangle T is put in W if its dual edge e is marked due to all three Voronoi cells defining e. Note that in the original algorithm [5], they estimate the surface

normal with the *polar normal*, which is the direction from p^* to the farthest vertex in the Voronoi cell V_{p^*} at p^*. Since we have already obtained good normal estimates ℓ_p^* in the POINT ESTIMATION step, we don't need to estimate the polar normals.

After the candidate triangles are collected in W, the algorithm proceed to the manifold extraction step. First, all triangles incident to *sharp edges* are deleted. An edge is called *sharp* if the angle between any two consecutive triangles around the edge is more than $3\pi/2$, or the edge is incident to only one triangle. Next, the outer boundary N of the remained triangles is extracted by a depth-first walk along the outer boundary of each of its connected components.

It is proved in [5] that N is a manifold homeomorphic to the surface S. In order to establish that, they need three basic important conditions, which are:

(I) restricted Delaunay condition: W contains all restricted triangles;

(II) small triangle condition: the circumradius of each triangle in W is $O(\epsilon)$; and

(III) flat triangle condition: the normal to each triangle $T \in W$ makes a small angle $O(\epsilon)$ with the surface normal at the vertex p where p is the vertex with the largest interior angle in T.

As in our case, all the center points in C^* may not lie on S, the proofs of the original cocone paper [5] do not automatically apply. We have to generalize them so that they apply to samples points not on surface S as well. In the following sections, we will first prove the three basic conditions. Note that the small value $O(\epsilon)$ will be replaced by $O(\epsilon^{\gamma/16})$. And then we go forward to prove the homeomorphism.

5.5.6 Restricted Delaunay condition

We show that all restricted Delaunay triangles belong to W. So W contains triangles that form a manifold homeomorphic to S by Lemma 5.5.6. This is essential as otherwise there is no hope that the cocone algorithm will succeed in extracting a manifold from W.

Lemma 5.5.7 *Let v be any restricted Voronoi vertex in $V_{p^*}|_S$. Then $\angle(p^*v, n_p) \geq \pi/2 - O(\epsilon^{\gamma/8})$.*

Proof. Let $p^*q^*s^*$ be the restricted Delaunay triangle dual to v. Let B be the Delaunay ball centered at v and passing through p^*, q^* and s^*. As $2d(p^*, v) = 2\text{radius}(B) \geq d(p^*, s^*) \geq \epsilon_2^{\gamma/8}$ by Lemma 5.5.3, we have $d(p^*, v) \geq d(p^*, s^*)/2 \geq \epsilon_2^{\gamma/8}/2$. On the other hand, as $v \in V_{p^*}|_S$, $d(p^*, v) = O(\epsilon^{\gamma/8})$ by Corollary 5.5.1. Also as $d(p^*, \tilde{p}) = O(\sqrt{r_p}) = O(\epsilon^{\gamma/2})$ by Lemma 5.4.6 and Lemma 5.3.1, we have $d(\tilde{p}, v) \leq d(p^*, \tilde{p}) + d(p^*, v) = O(\epsilon^{\gamma/8})$.

Now we have

$$\begin{aligned} \angle p^*v\tilde{p} &\leq \arcsin \frac{d(p^*, \tilde{p})}{d(p^*, v)} \\ &= \frac{O(\epsilon^{\gamma/2})}{\epsilon_2^{\gamma/8}/2} \\ &= \frac{O(\epsilon^{\gamma/2})}{\Omega(\epsilon^{(3/2)(\gamma/8)})} \\ &= O(\epsilon^{5\gamma/16}). \end{aligned}$$

By Lemma 4.1.2 (i), $\angle(\tilde{p}v, T_p) \leq \arcsin d(\tilde{p}, v) = O(\epsilon^{\gamma/8})$. So $\angle(\tilde{p}v, n_p) \geq \pi/2 - O(\epsilon^{\gamma/8})$. In all, $\angle(p^*v, n_p) \geq \angle(\tilde{p}v, n_p) - \angle p^*v\tilde{p} \geq \pi/2 - O(\epsilon^{\gamma/8}) - O(\epsilon^{5\gamma/16}) \geq \pi/2 - O(\epsilon^{\gamma/8})$. □

Lemma 5.5.8 *All restricted Delaunay triangles are in W.*

Proof. Let v be the restricted Voronoi vertex dual to the restricted Delaunay triangle $p^*q^*s^*$. By Lemma 5.5.7, $\angle(p^*v, n_p) \geq \pi/2 - O(\epsilon^{\gamma/8})$. Also $\angle(n_p, \ell_p^*) = O(\sqrt{r_p}) = O(\epsilon^{\gamma/2})$ by Lemma 5.4.6 and Lemma 5.3.1. The triangle inequality implies that $\angle(p^*v, \ell_p^*) \geq \angle(p^*v, n_p) - \angle(n_p, \ell_p^*) \geq \pi/2 - O(\epsilon^{\gamma/8}) - O(\epsilon^{\gamma/2}) \geq \pi/2 - O(\epsilon^{\gamma/8}) > \pi/2 - \pi/8$. Similarly, both $\angle(q^*v, n_q)$ and $\angle(s^*v, n_s)$ are greater than $\pi/2 - \pi/8$. Hence $p^*q^*s^*$ will be included in W by the algorithm. □

5.5.7 Small triangle condition

We prove that the circumcircle of each candidate triangle T in W is small; more precisely, its radius is at most $\epsilon^{\gamma/16}$.

Since there are twin big empty tangent balls tangent to S_α at center point $p^* \in S_\alpha$ by Lemma 5.5.4, we have the following lemma by using the same argument as in Lemma 5 in [2]. The details of this proof is included in the appendix.

Lemma 5.5.9 *Let v be any point in V_{p^*} such that $d(p^*, v) \geq \epsilon^{\gamma/16}$. Then $\angle(p^*v, n_{p^*}) = O(\epsilon^{\gamma/16})$.*

Lemma 5.5.10 *The radius of the smallest Delaunay ball of any triangle $T \in W$ is at most $\epsilon^{\gamma/16}$. Hence, the circumradius of T is at most $\epsilon^{\gamma/16}$.*

Proof. Let e be the dual edge of T, and p^* be a vertex of T. As the algorithm selects T, there is a point $v \in e$ so that $\angle(p^*v, \ell_p^*) \geq \pi/2 - \pi/8$. And by Lemma 5.4.6, $\angle(\ell_p^*, n_p) = O(\sqrt{r_p}) = O(\epsilon^{\gamma/2})$. Also by Lemma 5.5.1, $\angle(n_p, n_{p^*}) = O(\epsilon^{\gamma/2})$. Thus the triangle inequality implies that $\angle(p^*v, n_{p^*}) \geq \angle(p^*v, \ell_p^*) - \angle(\ell_p^*, n_p) - \angle(n_p, n_{p^*}) \geq \pi/2 - \pi/8 - O(\epsilon^{\gamma/2}) = \omega(\epsilon^{\gamma/16})$. Then by the contrapositive of Lemma 5.5.9, $d(p^*, v) < \epsilon^{\gamma/16}$. ∎

5.5.8 Flat triangle condition

We prove that the normal of each candidate triangle $T \in W$ makes a small angle of $O(\epsilon^{\gamma/16})$ radians with the surface normal at the vertex of T that has the largest vertex angle.

Since there are twin big empty balls tangent to S_α at center point $p^* \in S_\alpha$ by Lemma 5.5.4, and the candidate triangles in W are small by Lemma 5.5.10, we have the following theorem of triangle normal convergence by using the same argument in Theorem 11 in [5]. The details of this proof is in the appendix.

Lemma 5.5.11 *The normal to each candidate triangle $T \in W$ makes an angle of $O(\epsilon^{\gamma/16})$ radians with n_{p^*} where p^* is the vertex of T that has the largest vertex angle.*

This lemma says that the normal of triangle $T \in W$ is close to the surface normal at a vertex of T. As T is small by the small triangle condition, we can also show for any point x in the interior of T, $n_{\tilde{x}}$ is close to the normal of T.

Lemma 5.5.12 *Let x be any point on a triangle $T \in W$. Then*

(i) $d(x, \tilde{x}) = O(\epsilon^{\gamma/16})$, *and*

(ii) $\angle(n_{\tilde{x}}, n_T) = O(\epsilon^{\gamma/16})$, *where n_T is the normal line to T.*

Proof. Consider (i). Let p^* be a vertex of T. By Lemma 5.5.10, $d(x, p^*) \leq 2\epsilon^{\gamma/16}$, and by Lemma 5.4.6, $d(p^*, \tilde{p}) = O(\sqrt{r_p}) = O(\epsilon^{\gamma/2})$. So $d(x, \tilde{x}) \leq d(x, \tilde{p}) \leq d(x, p^*) + d(p^*, \tilde{p}) = O(\epsilon^{\gamma/16})$.

Consider (ii). By Lemma 5.5.1, $d(\tilde{p^*}, p^*) = O(\epsilon^{\gamma/2})$. By Lemma 5.5.10, $d(p^*, x) \leq 2\epsilon^{\gamma/16}$. By part (i), $d(x, \tilde{x}) = O(\epsilon^{\gamma/16})$. So $d(\tilde{p^*}, \tilde{x}) \leq d(\tilde{p^*}, p^*) + d(p^*, x) + d(x, \tilde{x}) = O(\epsilon^{\gamma/16})$. Thus by Lemma 4.1.2 (ii), $\angle(n_{p^*}, n_{\tilde{x}}) \leq 3d(\tilde{p^*}, \tilde{x}) = O(\epsilon^{\gamma/16})$. Also by Lemma 5.5.11, $\angle(n_T, n_{p^*}) = O(\epsilon^{\gamma/16})$. Hence $\angle(n_T, n_{\tilde{x}}) \leq \angle(n_T, n_{p^*}) + \angle(n_{p^*}, n_{\tilde{x}}) = O(\epsilon^{\gamma/16})$. ∎

5.5.9 Homeomorphism

A piecewise-linear manifold N is extracted from W in the manifold extraction step. It begins by recursively deleting any triangle in W adjacent to a sharp edge. Let W' be the remaining set of triangles. The lemma below proves that none of the restricted Delaunay triangles are deleted, and so W' contains a manifold homeomorphic to S.

Lemma 5.5.13 *No restricted Delaunay triangle has a sharp edge.*

Proof. First, since the restricted Delaunay triangles from a manifold homeomorphic to S, no edge of a restricted Delaunay triangle is incident to only one triangle. It remains to prove that no pair of restricted Delaunay triangles sharing an edge makes an acute angle.

Let T and T' be a pair of adjacent restricted Delaunay triangles sharing a common edge e. Let p^* be a vertex of e. Since T and T' are restricted Delaunay triangles, they have restricted Delaunay balls D, D' centered at restricted Voronoi vertices $v, v' \in S$ respectively.

Let H be the supporting plane of circle $\partial D \cap \partial D'$. Note that e lies on H and H separates T and T'. It is clear that $\angle(T, T') \geq \angle(T, H) + \angle(T', H)$ as H separates T and T'. So it suffices to lower bound $\angle(T, H)$ and $\angle(T', H)$.

Let n_T and $n_{T'}$ be the normal lines to T and T', respectively. By triangle inequality, $\angle(T, H) \geq \angle(n_p, T) - \angle(n_p, H)$. To lower bound $\angle(n_p, T)$, we need to upper bound $\angle(n_p, n_T)$. By Lemmas 5.5.1 and 5.5.5,

$$\angle(n_p, n_T) \leq \angle(n_{p^*}, n_p) + \angle(n_T, n_{p^*}) = O(\epsilon^{\gamma/2}) + O(\epsilon^{\gamma/8}) = O(\epsilon^{\gamma/8}).$$

So $\angle(n_p, T) \geq \pi/2 - O(\epsilon^{\gamma/8})$.

To upper bound $\angle(n_p, H)$, we need to lower bound $\angle(n_p, vv')$ as $vv' - H$. By Corollary 5.5.1 and Lemma 5.5.1, $d(v, \tilde{p}) \leq d(v, p^*) + d(p^*, \tilde{p}) = O(\epsilon^{\gamma/8}) + O(\epsilon^{\gamma/2}) = O(\epsilon^{\gamma/8})$. Then Lemma 4.1.2 (ii) implies that $\angle(n_v, n_p) \leq 3d(v, \tilde{p}) = O(\epsilon^{\gamma/8})$. So if we can lower bound $\angle(n_v, vv')$ then we are done. By Corollary 5.5.1, $d(v, v') \leq d(p^*, v) + d(p^*, v') = O(\epsilon^{\gamma/8})$. Then by Lemma 4.1.2 (i), $\angle(vv', T_v) \leq \arcsin d(v, v') = O(\epsilon^{\gamma/8})$. So $\angle(n_v, vv') = \pi/2 - \angle(vv', T_v) = \pi/2 - O(\epsilon^{\gamma/8})$. Hence

$$\angle(vv', n_p) \geq \angle(vv', n_v) - \angle(n_v, n_p) \geq \pi/2 - O(\epsilon^{\gamma/8}) - O(\epsilon^{\gamma/8}).$$

Thus $\angle(n_p, H) = \pi/2 - \angle(vv', n_p) = O(\epsilon^{\gamma/8})$, i.e., H is nearly parallel to n_p.

Then by triangle inequality, $\angle(T, H) \geq \angle(n_p, T) - \angle(n_p, H) \geq \pi/2 - O(\epsilon^{\gamma/8})$. Similarly we can prove that $\angle(T', H) \geq \pi/2 - O(\epsilon^{\gamma/8})$.

Now we have $\angle(T, T') \geq \angle(T, H) + \angle(T', H) \geq \pi - O(\epsilon^{\gamma/8})$, which is obtuse. So e is not sharp. ∎

Then in the algorithm, the outer boundary N of W' is extracted as the output manifold. In the lemma below, we prove that two adjacent triangles in N make an obtuse angle.

Lemma 5.5.14 *Every pair of adjacent triangles in N meets at their common vertex at an angle at least $\pi - O(\epsilon^{\gamma/16}) > \pi/2$.*

Proof. Suppose T and T' share a common vertex p^*. As $\angle(n_{p^*}, n_T) = O(\epsilon^{\gamma/16})$ and $\angle(n_{p^*}, n_{T'}) = O(\epsilon^{\gamma/16})$ by Lemma 5.5.11, we have

$$\angle(n_T, n_{T'}) \leq \angle(n_{p^*}, n_T) + \angle(n_{p^*}, n_{T'}) = O(\epsilon^{\gamma/16}).$$

So $\angle(T, T') \geq \pi - \angle(n_T, n_{T'}) \geq \pi - O(\epsilon^{\gamma/16}) > \pi/2$. ◻

Homeomorphism proof. Let $\mu : R^3 \to S$ be the map from each point in R^3 to the closest point on S. First we prove that the restriction of μ to N is well-defined.

Lemma 5.5.15 *$\mu : N \to S$ is well-defined.*

Proof. Suppose to the contrary that μ is not well-defined. Then there is a point x on a triangle in W such that x has more than one closest point to S. This means x is a point on the medial axis. So $d(x, S)$ is at least the local feature size of some point on S, and thus $d(x, S) \geq 1$. This contradicts Lemma 5.5.12(i) that $d(x, S) = O(\epsilon^{\gamma/16}) < 1$. ◻

Then we prove that $\mu : N \to S$ is a homeomorphism. A function is called a *homeomorphism* if it is continuous and bijective, and its inverse is also continuous. The approach is to first show that μ is well-behaved on center points C^*, and then extend the analysis to the interiors of the triangles in N.

Lemma 5.5.16 *The function $\mu : N \to S$ is continuous, and so is its inverse.*

Proof. By Lemma 5.5.10, every point $x \in N$ is within distance of $\epsilon^{\gamma/16}$ from a vertex $p^* \in C^*$, and thus is within a distance of $d(x,p^*) + d(p^*,\tilde{p}) = O(\epsilon^{\gamma/16})$ from $\tilde{p} \in S$ by Lemma 5.4.6. Note that the functions μ and μ^{-1} are continuous except at the medial axis of S. So the lemma follows from the fact that N and S are continuous and N and S avoid the medial axis. ◻

Lemma 5.5.17 *Assume $p^* \in C^*$ lies on S_α. Let m be the center of a medial ball M tangent to S_α at p^*. No candidate triangle in W intersects the interior of the segment p^*m.*

Proof. Let $M' = B(m', 1/10)$ be an empty ball tangent to S_α at p^* in Lemma 5.5.4. Note $m' \in p^*m$ as $d(p^*,m) = 1 - \Delta > 1/10$. Suppose to the contrary that a triangle T in W intersects the interior of p^*m at some point x. Then either $x \in m'm\backslash\{m\}$ or on $x \in \text{int}(p^*m')$. Note that $\tilde{x} = \tilde{p^*}$. By Lemma 5.5.1 and Lemma 5.5.12(i), $d(p^*,x) \leq d(p^*,\tilde{p^*}) + d(\tilde{x},x) = O(\epsilon^{\gamma/2}) + O(\epsilon^{\gamma/16}) = O(\epsilon^{\gamma/16})$.

If $x \in m'm\backslash\{m\}$, then $d(p^*,x) \geq 1/10 = \omega(\epsilon^{\gamma/16})$, which is a contradiction. Now suppose that $x \in \text{int}(p^*m')$. In order to intersect $\text{int}(p^*m')$, T has to intersect M' as $p^*m' \subset M'$, and so does the smallest Delaunay ball D of T. Let H be the supporting plane of the circle $\partial M' \cap \partial D$. By adapting the proof for Lemma 16 of [5], we can derive the contradiction that H separates $\text{int}(p^*m')$ and T.

On one side of H, M' is contained in D, and on the other, D is contained in M'. Let H^+ be the open halfspace, in which D contains M'. We first show that $p^* \notin H^+$. Since M' is empty of center points in C^*, T has to lie in H^+. Since D is Delaunay, p^* cannot lie in the interior of D. But since p^* lie on $\partial M'$, it therefore cannot lie in H^+. Next We show that $m' \notin H^+$ either. As $m' \in M'$, if $m' \in H^+$, m' has to lie in D. As $x \in T \subset D$, $xm' \subset D$. Since $2 \cdot \text{radius}(D) \geq d(x,m') \geq d(p^*,m') - d(p^*,x) \geq 1/10 - O(\epsilon^{\gamma/16})$, we have $\text{radius}(D) \geq 1/20 - O(\epsilon^{\gamma/16}) > \epsilon^{\gamma/16}$, which contradicts Lemma 5.5.10. In all, $p^*, m' \notin H^+$, which implies H separates $\text{int}(p^*m')$ and T. ◻

Corollary 5.5.2

(i) Every center point in C^* is a vertex of N.

(ii) The function μ is one-to-one from N to $\mu(p^*)$ for every $p^* \in C^*$.

Our homeomorphism proof proceeds in three short steps. We show that μ induces a homeomorphism on each triangle, then on each pair of adjacent triangles, and finally on N as a whole.

Now with Lemma 5.5.12(ii), by using the same arguments as in Lemma 18 in [5], we thus have the following lemma, which shows that μ induces a homeomorphism on each triangle, then on union of adjacent triangles around a vertex.

Lemma 5.5.18 *Let U be a region contained within one triangle $T \in N$ or in adjacent triangles of N. The function μ defines a homeomorphism between U and $\mu(U) \subset S$.*

Then by Corollary 5.5.2 and Lemma 5.5.18 and using same arguments as Theorem 19 in [5], we have that μ is a homeomorphism for the whole N.

Lemma 5.5.19 *The mapping μ defines a homeomorphism from the triangulation N to the surface S.*

Hence by the convergence lemmas (Lemmas 5.4.6, 5.5.11) , the homeomorphism lemma above (Lemma 5.5.19) and the lemma about the running time analysis (Lemma 5.1.2), we finally can get our main theorem for this chapter (Theorem 5.2.1).

5.6 Summary

The result in this paper successfully extends our result of curve reconstruction from noisy samples to surface reconstruction from noisy samples with theoretical guarantees. Instead of using probabilistic sampling model as in curve case, we use deterministic sampling model, which makes the proofs simpler and more intuitive. The algorithms runs in $O(n^{2+\gamma})$ time where $0 < \gamma \leq 1/8$ is a constant. The output

piecewise linear manifold is homeomorphic to S, is close to S in terms of Hausdorff distance, and has normals close to the surface normals on S. This result, together with the recent result by Dey and Goswami [15] are the first two results to do surface reconstruction from noisy samples with theoretical guarantees.

5.7 Appendix

Proof of Lemma 5.1.1

First, we claim that $\epsilon_2 = O(\sqrt[3]{\frac{\ln^\omega n}{n}})$. By this claim, since $\epsilon_2 = \Omega(\epsilon^{3/2})$ by the model, we have $\epsilon = O(\epsilon_2^{2/3}) = O((\frac{\ln^\omega n}{n})^{2/9})$. Then the upper bound on ϵ_1 follows as $\epsilon_1 = O(\epsilon)$ by the model.

Take a maximal set K of disjoint balls with radii $\epsilon_2/2$ and centers inside $\mathbf{S}$. So all points in K are within a distance of $\delta + \epsilon_2/2$ from S. A packing argument shows that $|K| \leq \frac{(2\delta+\epsilon_2)\text{area}(S)}{(4/3)\pi(\epsilon_2/2)^3} = O((\delta+\epsilon_2)/\epsilon_2^3)$. If we double the sizes of the balls in K, the expanded balls cover $\mathbf{S}$ entirely. Each expanded ball has radius ϵ_2 and contains at most $\ln^\omega n$ samples by the model. It follows that $|K| \geq n/\ln^\omega n$. Combining the upper and lower bounds for $|K|$, we get $n/\ln^\omega n = O((\delta+\epsilon_2)/\epsilon_2^3)$. If $\delta \leq \epsilon_2$, replacing δ by ϵ_2 yields $\epsilon_2 = O(\sqrt{\ln^\omega n/n})$. If $\delta > \epsilon_2$, as $\delta < 1$, replacing $\delta + \epsilon_2$ by 2 yields $\epsilon_2 = O(\sqrt[3]{\ln^\omega n/n})$. This proves our first claim.

Second, we claim that $\epsilon = \Omega(\sqrt{1/n})$. By this claim, since $\epsilon_2 = \Omega(\epsilon^{3/2})$ by the model, we have $\epsilon_2 = \Omega(1/n^{3/4})$.

Take a maximal set L of disjoint balls with radii ϵ and centered at points on S. Since each ball contains at least one sample by the model, $n \geq |L|$. By Lemma 4.3.3, the area of the intersection between S and each ball in L is $O(\epsilon^2)$. It follows that $|L| = \Omega(\text{area}(S)/\epsilon^2) = \Omega(1/\epsilon^2)$. Combining the upper and lower bounds for L, we get $\epsilon = \Omega(1/\sqrt{n})$. This proves our second claim. □

Proof of Lemma 5.3.3

Let ℓ be the support line of n_p. It suffices to show that as we translate ℓ on H to a distance d from p, $\ell \cap S_\alpha$ traces a single curve segment inside $B(p,d)$. By Lemma 4.4.2, $\ell \cap S_\alpha$ traces a single curve segment ξ inside $B(p,2d)$. Assume to the contrary that ξ exits $B(p,d)$ at a point x, and reenters $B(p,d)$ at a point y. Let $\xi(x,y)$ denote the subcurve between x and y that lies outside $B(p,d)$. By Lemma 4.1.2(i), $\angle(px, T_p) \leq \arcsin(d) \leq 2d$. Similarly, $\angle(py, T_p) \leq 2d$. Thus the

vector n from p to the midpoint of xy makes an angle at least $\pi/2 - 2d$ with n_p. Observe that n is parallel to a curve normal of some point $z \in \xi(x, y)$. It follows that $\angle(n_z, n_p) \geq \angle(n, n_p) \geq \pi/2 - 2d$. But this contradicts Lemma 4.1.2(ii). □

Proof of Lemma 5.4.1

Let $C = L_p(\ell) \cap S_\sigma \cap \mathit{coarse}(p)$. It follows from Lemma 4.4.2 that C is a topological disk, and for any line $h \in L_p(\ell)$, since $d(p, h) \leq r_p/4$, h intersects $S_\sigma \cap B(x, r_p/2)$ exactly once. Hence $C \subseteq B(x, r_p/2)$. □

We then prove Lemma 5.4.2, which bounds some distances from p to S_α in terms of $\angle(n_p, \ell)$.

Proof of Lemma 5.4.2

Refer to Figure 5.6. Let h be the line through p_1 parallel to ℓ. Let Y be the infinite tube with axis h and p on its boundary. So the radius of Y is $d(p, p_1) \sin\theta$. Then as in the proof of Lemma 5.4.1, $d(p_1, x) \leq \text{radius}(Y)/2$, which is at most $d(p, p_1)/2 \cdot \sin\theta \leq \theta d(p, p_1)/2$.

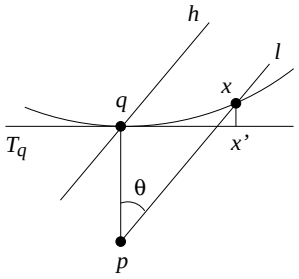

Figure 5.6: Illustration for the proof of Lemma 5.4.2.

Let x' be the orthogonal projection of x onto T_{p_1}. By Lemma 4.1.2(i), $d(x, x') \leq d(p_1, x)^2$, which is at most $\theta^2 d(p, p_1)^2/4$. Thus

$$\begin{aligned} d(p, x) &\leq \frac{d(p, p_1) + d(x, x')}{\cos\theta} \\ &\leq \frac{d(p, p_1) + \theta^2 d(p, p_1)^2/4}{\cos\theta}. \end{aligned}$$

Symmetrically, $d(p, x) \geq \frac{d(p,p_1) - d(x,x')}{\cos\theta}$ gives the lower bound. □

Proof of Lemma 5.5.5

Let s^* be the vertex of τ that has the largest vertex angle ($\geq \pi/3$). Suppose that s^* is on the surface S_α. Let C be the circumcircle of T. By Lemma 5.5.4, there are twin empty balls B, B' of radius 1/10 tangent to S_α at s^*. These balls intersect the supporting plane of τ in twin disks D, D' tangent at s^* such that $D \subset B$ and $D' \subset B'$. Let R be the common radius of D and D'. See Figure 5.7. Our first upper bound R in terms of radius(C).

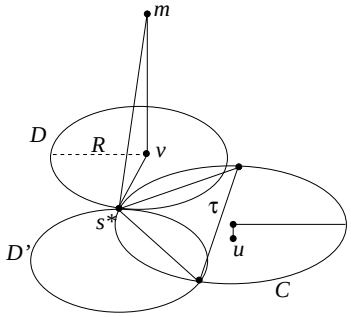

Figure 5.7: Illustration for the proof of Lemma 5.5.5.

Since the twin tangent balls at s^* are empty, the twin disks do not contain the vertices of τ. In order to maximize R relative to radius(C), we assume that the twin disks pass through the vertices of τ and that the angle at s^* measures exactly $\pi/3$. Now it is not hard to show that R is maximized exactly when T is equilateral: if we move s^* away from the midpoint of the arc covered by the twin disks, keeping the twin disks passing through the vertices of T, the radius R decreases, until s^* reaches one of the other vertices of τ and $R =$ radius(C). Since the worst-case configuration is equilateral T, we can conclude that $R \leq \sqrt{3}$radius(C).

Now we can bound all these radii. Let u denote the restricted Voronoi vertex dual to T. Since u is the minimum circumradius of τ, radius(C) $\leq d(u, s^*)$. By Corollary 5.5.1, $d(u, s^*) = O(\epsilon^{\gamma/8})$, which implies that $R = O(\text{radius}(C)) = O(\epsilon^{\gamma/8})$.

Now to find the angle between the normal to τ and the normal to S_α at s^*, we consider one of the twin tangent balls B at s^*. Let m denote the center of B and let v denote the center of D. The segment s^*m is normal to S_α at s^* and the segment mv is normal to T, so the angle $\angle(n_\tau, n_{s^*})$ we would like to bound is

$\angle s^*mv$. Observe that $\angle s^*mv = \pi/2$, $d(s^*, m) = 1/10$, and $d(s^*, v) = R = O(\epsilon^{\gamma/8})$. Hence $\angle(n_\tau, n_{s^*}) = \angle s^*mv \leq \arcsin(O(\epsilon^{\gamma/8})) = O(\epsilon^{\gamma/8})$.

Take any vertex p^* of τ other than s^*. By Corollary 5.5.1, $d(u, s^*) = O(\epsilon^{\gamma/8})$ and $d(u, p^*) = O(\epsilon^{\gamma/8})$. Thus $d(s^*, p^*) \leq d(u, s^*) + d(u, p^*) = O(\epsilon^{\gamma/8})$. Then

$$\begin{aligned} d(\tilde{s^*}, \tilde{p^*}) &\leq d(s^*, p^*) + d(s^*, \tilde{s^*}) + d(p^*, \tilde{p^*}) \\ &= O(\epsilon^{\gamma/8}) + 2\Delta \\ &= O(\epsilon^{\gamma/8}) + O(\epsilon^{\gamma/2}) \\ &= O(\epsilon^{\gamma/8}). \end{aligned}$$

And by Lemma 4.1.2(ii), we have $\angle(n_{s^*}, n_{p^*}) \leq 3d(\tilde{s^*}, \tilde{p^*}) = O(\epsilon^{\gamma/8})$. Hence $\angle(n_T, n_{p^*}) \leq \angle(n_T, n_{s^*}) + \angle(n_{s^*}, n_{p^*}) = O(\epsilon^{\gamma/8})$. □

To prove Lemma 5.5.6, we need Lemma 5.5.5 and two more lemmas from paper [2] as stated below.

Lemma 5.7.1 *For any two points p and q on S with $d(p, q) \leq \lambda f(p)$, $\angle(pq, n_p) \geq \pi/2 - \arcsin(\lambda/2)$.*

Lemma 5.7.2 *If a ball B intersects surface S in more than one connected component, then B contains a point of the medial axis of S.*

With these lemmas, by using almost the same arguments as Theorem 2 in [2], we can prove that the closed ball property: the closure of each k-dimensional face, $1 \leq k \leq 3$, of the Voronoi diagram of C^* intersects intersect S in either the empty set *or* in a closed $(k-1)$-dimensional topological ball. With this property, by the theorem (Theorem 4.3) of Edelsbrunner and Shah [23], we immediately have Lemma 5.5.6.

Proof of Lemma 5.5.6

To prove the closed ball property, we will consider the intersection of a Voronoi edge, a Voronoi face, and a Voronoi cell with the surface S respectively.

Let p^* be a selected center point in C^*, and let V_{p^*} be its Voronoi cell. Let the direction of n_{p^*} be vertical.

We begin by showing that in the vicinity of p^*, the surface S is nearly horizontal. Corollary 5.5.1 shows that $V_{p^*}|_S$ is small, fitting inside a ball B centered at p^* of radius $O(\epsilon^{\gamma/8})$. Since such a small ball cannot intersect the medial axis, Lemma 5.7.2 implies that $S \cap B$ is connected. Lemma 4.1.2(ii) shows that the normal to $S \cap B$ makes an angle $O(\epsilon^{\gamma/8})$ radians with the vertical.

Consider an edge e of V_{p^*}. If e intersects S, then since e is normal to the dual Delaunay triangle T, Lemma 5.5.5 implies that e is within $O(\epsilon^{\gamma/8})$ radians from the normal to S at p^*. So e can intersect S only once within ball B.

Next consider a face f of V_{p^*}. The face f is contained in its supporting plane H, the perpendicular bisector of p^* and another center point s^*, where p^*s^* is an edge of a restricted triangle T. The plane H must contain a vector h parallel to the normal of T, so again Lemma 5.5.5 shows that the angle $\angle(h, n_{p^*})$ is at most $O(\epsilon^{\gamma/8})$ radians. Each component of $f \cap S$ is an arc of a curve, with endpoints on the edges of f. Assume for the sake of contradiction that there are at least two such connected components, and consider any line segment connecting a point on one component with a point on another. Since $V_{p^*}|_S$ is small, Lemma 5.7.1 implies that each of these line segment is *nearly horizontal*, specifically within angle $O(\epsilon^{\gamma/8})$ radians with the horizontal. Hence we can sort the arcs of $f \cap S$ from left to right across f, as shown in Figure 5.7(a). Let q_1q_2 be a line segment connecting the right endpoint of one arc with the left endpoint of the next arc. The segment q_1q_2 is nearly horizontal so it must leave f as it crosses the nearly vertical edge of V_{p^*} at q_1 and reenter f at q_2, a contradiction to the fact that f is convex.

Finally consider V_{p^*} itself. Let C be $V_{p^*}|_S$. Now C cannot have a handle because S is nearly horizontal everywhere within ball B. We assume, however, that C is not a topological disk, again aiming for a contradiction. If C has no handles and is not a topological disk, then either it has more than one connected component or it is a topological disk with holes.

Consider the projection C' of C onto a horizontal plane. Since each pair of points in C are connected by a nearly horizontal segment, this projection is one-to-one, and C' is a planar shape homeomorphic to C. If C' has more than one connected component, let σ be the shortest segment connecting two different components, and

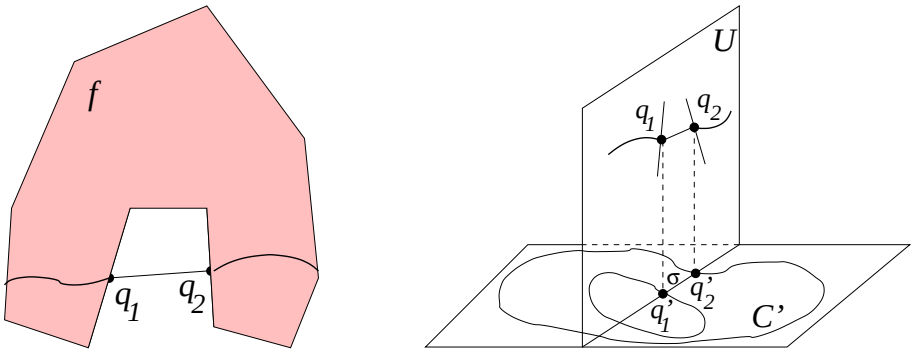

Figure 5.8: Two impossible situations: (a) Face f of V_{p^*} intersects S in two arcs. (b) $V_{p^*} \cap S$ is a disk with a hole.

let q_1', q_2' be that endpoints of σ. Otherwise, C' has a hole. Select some point x in a hole, let σ' be the line segment connecting x to the nearest point q_1' in C'. The line supporting σ' intersects C' again on the other side of x in a point q_2'. We let σ be the segment $q_1'q_2'$. Please refer to Figure 5.7(b).

Segment σ is perpendicular to the boundary of C' at q_1'. Let U be the vertical plane through σ, and let q_1, q_2 be the points in $U \cap C$ whose vertical projections are q_1', q_2' respectively. The point q_1 is either a vertex of C or an edge point, and thus lies in either an edge or facet of V_{p^*} respectively.

Consider the case in which q_1 lies in a facet f of V_{p^*}. Facet f is nearly vertical, and since U is perpendicular to C' at q_1', it is nearly perpendicular to C at q_1, and f intersects U in a nearly vertical line at q_1. In the other case, q_1 is contained in a nearly vertical edge e of V_{p^*}. We consider the plane V containing both e and the horizontal line perpendicular to segment q_1q_2 at q_1. V meets U in a nearly vertical line. Note that we had to choose U carefully, since two nearly vertical planes which are not nearly perpendicular might meet in a nearly horizontal line.

In either case, examining the situation in U, we find that in the neighborhood of q_1, the interior of q_1q_2 is separated from the interior of $V_{p^*} \cap U$ by a nearly vertical line. However, both q_1 and q_2 belong to $V_{p^*} \cap U$, contradicting to the fact that V_{p^*} is convex. □

Proof of Lemma 5.5.9

Suppose p^* is on the surface S_α, where $\alpha \leq \Delta$. Let B be the Voronoi ball centered at v. Let $M_1 = B(m_1, 1/10), M_2 = B(m_2, 1/10)$ be the twin empty balls tangent to S_α at p^* in Lemma 5.5.4. Suppose v and m_1 lie on the same side of S_α. Refer to Figure 5.9.

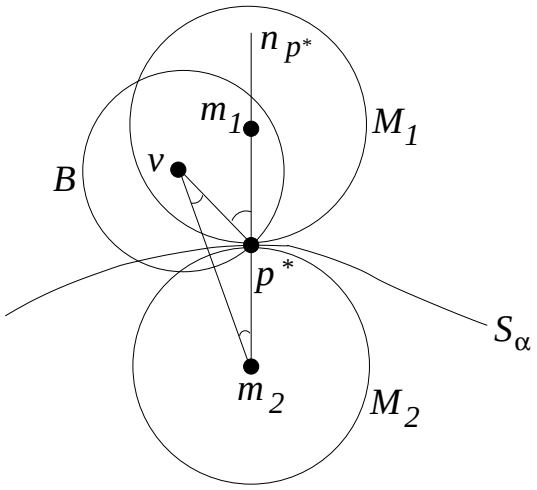

Figure 5.9: Illustration for the proof of Lemma 5.5.9.

As m_2 and v lie on opposite sides of S_α, line segment m_2v must intersect S_α at a point, say q. As M_1 and M_2 are empty, m_1 and m_2 belong to V_{p^*}. So $m_2v \subset V_{p^*}$, which implies $q \in V_{p^*}$. By Corollary 5.5.1, $d(q, p^*) = O(\epsilon^{\gamma/8})$. It is clear that $d(q, p^*) \geq d(v, p^*) \sin \angle v \geq \epsilon^{\gamma/16} \sin \angle v$. Thus $\angle v = O(\epsilon^{\gamma/8})/\epsilon^{\gamma/16} = O(\epsilon^{\gamma/16})$. On the other hand, $d(q, p^*) \geq d(v, p^*) \sin \angle m_2 = (\sin \angle m_2)/10$. Thus $\angle m_2 = O(\epsilon^{\gamma/8})$. Hence $\angle(p^*v, n_{p^*}) = \angle vp^*m_1 \leq \angle m_2 + \angle v = O(\epsilon^{\gamma/16}) + O(\epsilon^{\gamma/8}) = O(\epsilon^{\gamma/16})$. □

Proof of Lemma 5.5.11

Suppose $p^* \in S_\alpha$. Consider the twin empty balls M_1 and M_2 touching S_α at p^* of radius 1/10. Let D be the ball with the circumcircle of T as a diameter; refer to Figure 5.10(a). The radius r of D is equal to the radius of the circumcircle of T. Denote the circles of the intersection of D with M_1 and M_2 as C_1 and C_2 respectively. The normal to S at p passes through m, the center of M_1. This normal makes an

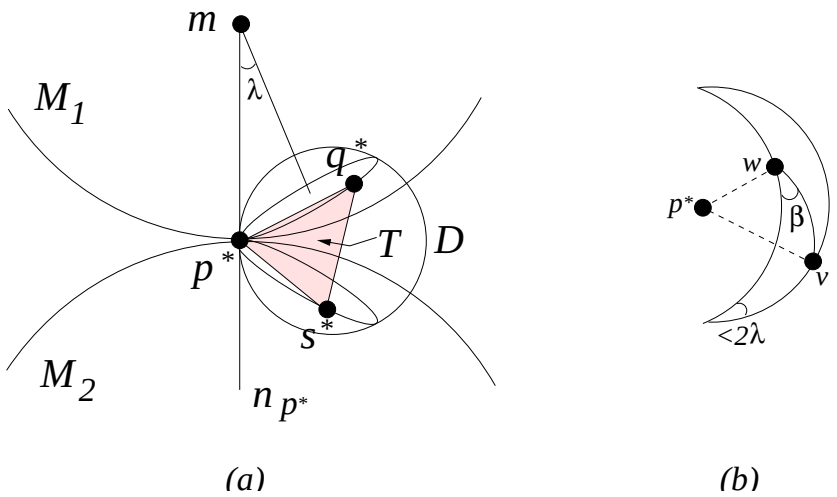

Figure 5.10: Normal of a small triangle T and the normal to S_α at the vertex p^* of T with the largest vertex angle.

angle less than λ with the normals to the planes of C_1 and C_2, where

$$\begin{aligned} \lambda &\leq \arcsin(r/d(p^*, m)) \\ &= \arcsin(10r) \\ &= O(\epsilon^{\gamma/16}) \end{aligned}$$

by Lemma 5.5.10. This angle bound also applies to the plane of C_2, which implies that the planes of C_1 and C_2 make a wedge, say G, with an acute angle no more than 2λ.

The other two vertices q^* and s^* of T cannot lie inside M_1 and M_2. This implies that T lies completely in the wedge G. Consider a cone at p^* inside the wedge G formed by the three planes: the plane π_T supporting T, the plane π_1 supporting C_1, and the plane π_2 supporting C_2. A unit sphere centered around p^* intersects the cone in a spherical triangle uvw, where u, v and w are the intersection points of the lines $\pi_1 \cap \pi_2$, $\pi_T \cap \pi_1$ and $\pi_T \cap \pi_2$ with the unit sphere respectively. See Figure 5.10(b). Without the loss of generality, assume that the angle $\angle uvw \leq \angle uwv$. We have the following facts. The arc length of wv, denoted by $|wv|$, is at least $\pi/3$ since p^* subtends the largest angle in T and T lies completely in the wedge G. The spherical angle $\angle vuw$ is less than or equal to 2λ. We are interested in the spherical angle $\beta = \angle uvw$ which is also the acute dihedral angle between the planes of T and C_1. By standard *sine laws* in spherical geometry, we have $\sin\beta =$

$\sin|uw|\frac{\sin\angle vuw}{\sin|wv|} \leq \sin\sin|uw|\frac{\sin(2\lambda)}{\sin|wv|}$. If $\pi/3 \leq |wv| \leq 2\pi/3$, we have $\sin\pi/3 \geq \sqrt{3}/2$ and hence $\beta \leq \arcsin\frac{2}{\sqrt{3}}\sin(2\lambda)$. For the range $2\pi/3 < |wv| \leq \pi$, we use the fact that $|uw|+|wv| \leq \pi$ since $\angle vuw \leq 2\alpha < \pi/2$ for sufficiently small ϵ. So, in this case $\frac{\sin|uw|}{\sin|wv|} < 1$. Thus $\beta \leq \arcsin\frac{2}{\sqrt{3}}\sin(2\lambda) = O(\epsilon^{\gamma/16})$.

Hence the angle between the normal of T and n_{p^*} is an acute angle at most $\lambda+\beta = O(\epsilon^{\gamma/16})$. □

Proof of Lemma 5.5.18

We know that μ is well-defined and continuous on U, and its inverse is also continuous, so it only remains to show that it is one-to-one. First, we prove that if U is in one triangle T, μ is one-to-one. For a point $q \in T$, the vector from q to $\mu(q)$ is normal to the surface at $\mu(q)$. If there was some $y \neq q$ on T with $\mu(y) = \mu(q)$, then q, $\mu(q)$ and y would all be collinear and T itself would have to contain line segment qy, contradicting Lemma 5.5.12(ii), which says that the normal of T is nearly parallel to the normal of S at $\mu(q)$.

Now, we consider the case in which U is contained in more than one triangle. Let q and y be two points U such that $\mu(q) = \mu(y) = x$, and let p^* be a common vertex of the triangles that contain U. Since μ is one-to-one in one triangle, q and y must lie in the two distinct triangles T_q and T_y respectively. Let l be the normal line to S at x. Note that $l \parallel n_x$. l pieces the patch U at least twice; if y and q are not adjacent intersections along l, redefine q so that this is true. Now consider the orientation of the patch U according to the estimated normal direction ℓ_p^* at p^*. Either l passes from inside to outside and back to inside when crossing y and q, or from outside to inside and back to outside.

The acute angles between the triangle normals of T_q, T_y and n_x are $O(\epsilon^{\gamma/16})$ by Lemma 5.5.12(ii), that is, the triangles are stabbed nearly perpendicularly by n_x. But since the orientation of U is opposite at the two intersections, the angle between the two *oriented* triangle normals is greater than zero, meaning that T_q and T_y must meet at p^* at an acute angle. This would contradict Lemma 5.5.14, which is that T_q and T_y meet at p^* at an obtuse angle. Hence there are no two points y, q in U

with $\mu(q) = \mu(y)$. ◻

Proof of Lemma 5.5.19

Let $S' \subset S$ be $\mu(N)$. We first show that (N, μ) is a *covering space* of S'. Informally, (N, μ) is a covering space for S' if function μ maps N smoothly onto S', with no folds or other singularities; see Massey [41], Chapter 5. Showing that (N, μ) is a covering space is weaker than showing that μ defines a homeomorphism, since, for instance, it does not preclude several connected components of N mapping onto the same component of S', or more interesting behavior, such as torus wrapping twice around another torus wrapping twice around another torus to form a *double covering*.

Formally, the (N, μ) is a covering space of S' if, for every $x \in S'$, there is a path-connected *elementary neighborhood* V_x around x such that each path-connected component of $\mu^{-1}(V_x)$ is mapped homeomorphically onto V_x by μ.

To construct such an elementary neighborhood, note that the set of points $|\mu^{-1}(x)|$ corresponding to a point $x \in S'$ is non-zero and finite, since μ is one-to-one on each triangle of N and there are only a finite number of triangles. For each point $q \in \mu^{-1}(x)$, we choose an open neighborhood U_q of around q, homeomorphic to a disk and small enough so that U_q is contained only in triangles that contain q. Refer to Figure 5.11.

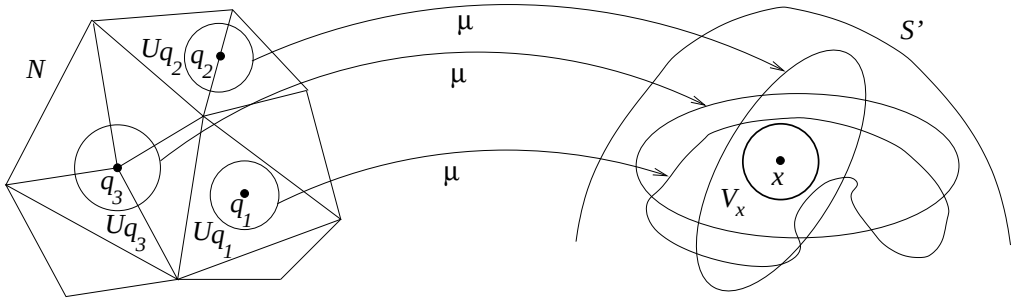

Figure 5.11: Illustration for the proof of Lemma 5.5.19.

As m_2 and v lie on opposite sides of S_α,

We claim that μ maps each U_q homeomorphically onto $\mu(U_q)$. This is because it is continuous, it is onto $\mu(U_q)$ by definition, and, since any two points x and y in

U_q are in adjacent triangles, it is one-to-one by Lemma 5.5.18.

Let $U'(x) = \cap_{q \in \mu^{-1}(x)} \mu(U_q)$, the intersection of the maps of each of the U_q. $U'(x)$ is the intersection of a finite number of open neighborhoods, each containing x, so we can find an open disk V_x around x. V_x is path connected, and each component of $\mu^{-1}(V_x)$ is a subset of some U_q and hence is mapped homeomorphically onto V_x by μ. Thus (N, μ) is a covering space for S'.

We show that μ defines a homeomorphism between N and S'. Since N is onto S' by definition, we need only show that μ is ono-to-one. Consider one connected component G of S'. A theorem of algebraic topology (see eg. Massey [41], Chapter 5, Lemma 3.4) says that when (N, μ) is a covering space of S', the sets $\mu^{-1}(x)$ for all $x \in G$ have the same cardinality. We now use Corollary 5.5.2(ii), that μ is one-to-one from N to $\mu(p^*)$ for every $p^* \in C^*$. From the algorithm, it is clear that each connected component of S contains a $\mu(p^*)$ for some center point $p^* \in C^*$. So it must be the case that $\mu : N \to S'$ is everywhere one-to-one, and N and S' are homeomorphic.

Finally, we show that $S' = S$. Since N is closed and compact, S' must be as well. So S' cannot include part of a connected component of S, and hence S' must consist of a subset of the connected components of S. Since every connected component of S contains a $\mu(p^*)$ for some center point $p^* \in C^*$, all components of S belong to S'. Hence $S' = S$, and N and S are homeomorphic. ◻

CHAPTER 6

CONCLUSION AND DISCUSSION

In the literature, most of the faithful reconstruction algorithms are proposed for noiseless sample points from smooth curves / surfaces, or from curves with sharp corners and endpoints. In this thesis, we propose provably faithful reconstruction algorithms for noisy sample points lying close to the original curve / surface.

We first propose a probabilistic noise model for the curve reconstruction problem. Based on this model, we design a curve reconstruction algorithm for the noisy input. The reconstruction is faithful with probability approaching 1 as the sampling density increases. Then we extend our approach to perform surface reconstruction from the noisy input points. We are able to make the noise model deterministic so that the analysis of the faithfulness proof is simplified. We prove that the surface reconstructed is faithful to the original surface if the input points satisfy the deterministic noise model. Both of our curve and surface reconstruction algorithms follow the same framework. First observe that all the sampling points inside the coarse ball centering at any sample point p fall inside a thin slab. The normal to this thin slab is a rough estimate to the surface normal. However, the deviated angle between this rough normal and the real surface normal depends on the noise amplitude δ, which is a constant by our assumption. That means the deviated angle does not tend to zero when the sampling density increases. Therefore, the algorithms move on to another stage to estimate a finer normal direction in the so-called refined neighborhood. Using this refined normal, the position of the sample point p is re-estimated to a new point (called center point), which is closer to the original curve / surface than p itself. A subset of these center points is then selected for the final surface reconstruction, which can be done using any existing faithful reconstruction algorithm for noiseless samples.

Dey & Goswami [15] recently has also proposed a provable reconstruction algorithm for noisy sample points lying close to a smooth surface. But in their model,

the noise amplitude is inversely proportional to the sampling density. That means they assume the sample points converges to the original manifold automatically when the sampling density increases. Instead, we assume that the noise amplitude is a constant globally. In order to have the point convergence property, we have to re-estimate the sample point positions to get a set of center points converging to the original manifold. Note that one more difference is that in their noise model, they assume the noise amplitude varies along the manifold and is proportional to the local feature size. One advantage of this assumption is that at the area where the local feature size is large, a large noise amplitude is allowed. One immediate future work would be to generalize our faithful reconstruction algorithm for the sample points satisfying a more general noise model such that the noise amplitude is a constant fraction of the local feature size.

Currently, our assumption that the noises are close to the original manifold are quite restrictive. We are investigating whether this approach is helpful to handle outlier noise. Usually the far-away outlier points are much more sparse than the sample points closer to the original manifold. That means that the majority of the sample points cluster around the original manifold, but only a small proportion of them scatters around. See Figure 6.1. In such case, it seems that in the coarse

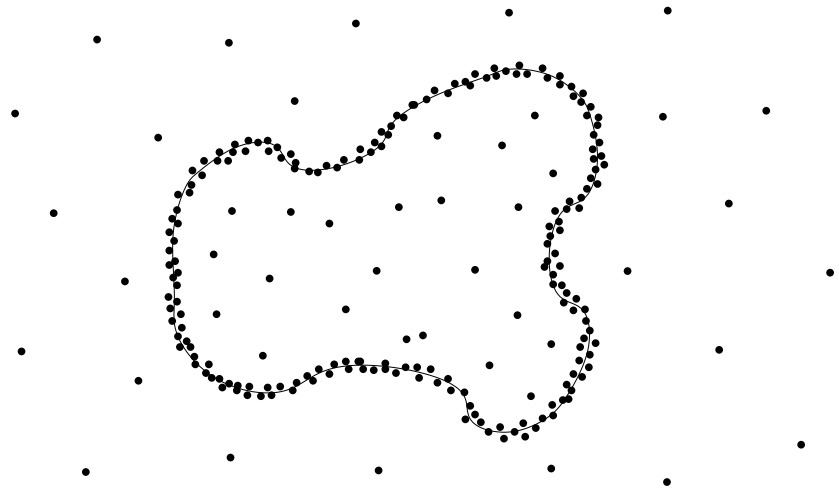

Figure 6.1: Noisy sample points with outliers.

neighborhood, a thin slab still can be formed by only enclosing a large proportion of sample points inside the whole coarse ball. Similarly in the refined neighborhood,

we can consider the min-height cylinder which encloses a large proportion of sample points inside the cylinder and the corresponding coarse ball. This cylinder direction should also approximate well to the surface normal. See Figure 6.2.

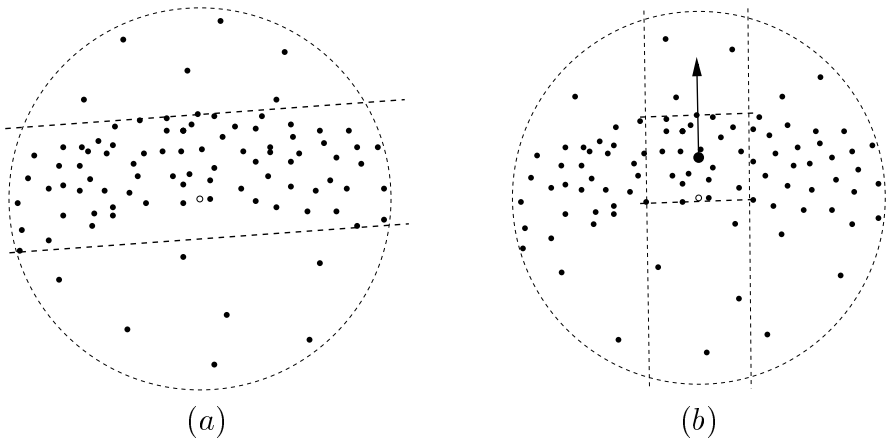

Figure 6.2: (a) Coarse neighborhood. (b) Refined neighborhood.

We also suspect that our approach sheds some light on reconstructing curves and surfaces with features such as corners, branchings, endpoints and noise altogether. These problems awaits to be answered.

REFERENCES

[1] E. Althaus and K. Mehlhorn. Traveling salesman based curve reconstruction in polynomial time. *SIAM Journal on Computing*, 31, 2001, 27–66.

[2] N. Amenta and M. Bern. Surface reconstruction by Voronoi filtering. *Discrete and Computational Geometry*, 22, 1999, 481–504.

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

[4] N. Amenta, M Bern and M. Kamvysselis. A new Voronoi-based surface reconstruction algorithm. *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, 1998, 415–421.

[5] N. Amenta, S. Choi, T. Dey and N. Leekha. A simple algorithm for homeomorphic surface reconstruction. *International Journal on Computational Geometry & Applications*, 12, 2002, 125–141.

[6] N. Amenta, S. Choi and R.K. Kolluri. The power crust. *Proceedings of the 6th ACM Symposium on Solid Modeling and Applications*, 2001, 249–260.

[7] N. Amenta, S. Choi and R.K. Kolluri. The power crust, unions of balls, and the medial axis transform. *Computational Geometry: Theory and Applications*, 19, 2001, 127–153.

[8] J.-D. Boissonnat and F. Cazals. Smooth surface reconstruction via natural neighbour interpolation of distance functions. *Computational Geometry*, 22, 2002, 185–203.

[9] R. Boyle and R. Thomas. *Computer Vision: A First Course*, Blackwell Scientific Publications, 1988, 32–34.

[10] J. Brandt and V.R. Algazi. Continuous skeleton computation by Voronoi diagram. *Computer Vision, Graphics and Image Processing*, 55, 1992, 274–280.

[11] S.-W. Cheng, S. Funke, M. Golin, P. Kumar, S.-H. Poon, and E.A. Ramos. Curve reconstruction from noisy samples. *Proceedings of the 19th Annual Symposium on Computational Geometry*, 2003, 302–311.

[12] B. Curless and M. Levoy. A volumetric method for building complex models from range images. *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, 1996, 303–312.

[13] J.-P. Dedieu and Ch. Favardin. Algorithms for ordering unorganized points along parametrized curves. *Numerical Algorithms 6*, 1994, 169–200.

[14] T.K. Dey, J. Giesen, S. Goswami, J. Hudson, R. Wenger, and W. Zhao. Undersampling and oversampling in sample based shape modeling. *Proceedings of IEEE Visualization*, 2001, 83–90.

[15] T.K. Dey and S. Goswami. Provable surface reconstruction from noisy samples. *Proceedings of the 20th Annual Symposium on Computational Geometry*, 2004, 330–339.

[16] T.K. Dey and P. Kumar. A simple provable algorithm for curve reconstruction. *Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms*, 1999, 893–894.

[17] T.K. Dey, K. Mehlhorn and E. Ramos. Curve reconstruction: connecting dots with good reason. *Computational Geometry: Theory & Applications*, 15, 2000, 229–244.

[18] T.K. Dey and R. Wenger. Reconstructing curves with sharp corners. *Computational Geometry: Theory & Applications*, 19, 2001, 89–99.

[19] T.K. Dey and R. Wenger. Fast reconstruction of curves with sharp corners. *Computational Geometry: Theory & Applications*, 12, 2002, 353–400.

[20] H. Edelsbrunner. Surface reconstruction by wrapping finite point sets in space. *Ricky Pollack and Eli Goodman Festschrift*, ed. B. Aronov, S. Basu, J. Pach and M. Sharir, Springer-Verlag, 2003, 379-404.

[21] H. Edelsbrunner, D.G. Kirkpatrick, and R. Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29, 1983, 551–559.

[22] H. Edelsbrunner and E.P. Mucke. Three-dimensional alpha shapes. *ACM Transactions on Graphics*, 13, 1994, 43-72.

[23] H. Edelsbrunner and N. Shah. Triangulating topological spaces. *International Journal on Computational Geometry and Applications*, 7, 1997, 365–378.

[24] L. Fang and D.C. Gossard. Fitting 3D curves to unorganized data points using deformable curves. *Computer Graphics International*, 1992, 535–543.

[25] S. Funke and E.A. Ramos. Reconstructing a collection of curves with corners and endpoints. *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithm*, 2001, 344–353.

[26] S. Funke and E. A. Ramos. Smooth-surface reconstruction in near-linear time. *Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithm*, 2002, 781–790.

[27] G. Gerig, O. Kubler, R. Kikinis, and F.A. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11, 1992, 221–232.

[28] J. Giesen. Curve reconstruction, the traveling salesman problem and Menger's theorem on length. *Discrete & Computational Geometry*, 24, 2000, 577–603.

[29] J. Giesen and U. Wagner. Shape dimension and intrinsic metric from samples of manifolds. *Proceedings of the 19th Annual ACM Symposium on Computational Geometry*, 2003, 329–337.

[30] C. Gold and J. Snoeyink. A one-step crust and skeleton extraction algorithm. *Algorithmica*, 30, 2001, 144–163.

[31] A.A. Goshtasby. Grouping and parameterizing irregularly spaced points for curve fitting. *ACM Transactions on Graphics*, 19, 2000, 185–203.

[32] M. Habib, C. McDiarmid, J. Ramirez-Alfonsin and B. Reed. *Probabilistic Methods for Algorithmic Discrete Mathematics*, Springer Verlag Belin Heidelberg, 1998, 198-200.

[33] S. Har-Peled and K.R. Varadarajan. Approximate shape fitting via linearization. *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, 2001, 66–73.

[34] H. Hoppe. Surface reconstruction from unorganized points. *Ph.D. Thesis*, Computer Science and Engineering, University of Washington, 1994.

[35] H. Hoppe, T. DeRose, T. Duchamp, H. Jin, J. McDonald, and W. Stuetzle. Piecewise smooth surface reconstruction. *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*, 1994, 19–26.

[36] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Surface reconstruction from unorganized points. *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques*, 1992, 71–78.

[37] K.I. Lee. Curve reconstruction from unorganized points. *Computer Aided Geometric Design*, 17, 2000, 161–177.

[38] D. Levin. The approximation power of moving least-squares. *Mathematics of Computation*, 67, 1998, 1517–1531.

[39] D. Levin. Mesh-independent surface interpolation. *Geometric Modeling for Scientific Visualization* (eds. Brunnett, Hamann and Mueller), Springer-Verlag, 2003.

[40] G. Medioni and C.K. Tang. Inference of integrated surface, curve, and junction descriptions from sparse 3-D data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1998, 1206–1223.

[41] W.S. Massey. *Algebraic Topology: An Introduction*, Springer-Verlag, Graduate texts in Mathematics 56, 1967.

[42] N.J. Mitra and A. Nguyen. Estimating surface normals in noisy point cloud data. *Proceedings of the 14th Annual ACM Symposium on Computational Geometry*, 2003, 322–328.

[43] H. Pottmann and T. Randrup. Rotational and helical surface approximation for reverse engineering. *Computing*, 60, 1998, 307–322.

[44] J. O'Rourke. *Computational Geometry in C*, Cambridge University Press, 1994.

[45] J. O'Rourke, H. Booth, R. Washington. Connect-the-dots: a new heuristic. *Computer Vision, Graphics and Image Processing*, 39, 1984, 258–266.

[46] G.P. Robinson, A.C.F. Colchester, L.D. Griffin, and D.J. Hawkes. Integrated skeleton and boundary shape representation for medical image interpretation. *Proceedings of the European Conference on Computer Vision*, 1992, 725–729.

[47] P. Saint-Marc, J.S. Chen, and G. Medioni. Adaptive smoothing: a general tool for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 1990, 629–639.

[48] J. Sijbers, A.J. den Dekker, A. Van der Linden, M. Verhoye, and D. Van Dyck. Adaptive anisotropic noise filtering for magnitude MR data. *Magnetica Resonance Imaging*, 17, 1999, 1533–1539.

[49] C.K. Tang and G. Medioni. A curvature-augmented tensorial framework for shape inference from noisy, 3-D data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 2002, 858–864.

[50] G. Taubin and R. Ronfard. Implicit simplicial models for adaptive curve reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18, 1996, 321–325.

[51] J.W. Turkey. *Exploratory Data Analysis*, John Wiley and Sons, Inc, 1991.

[52] Y. Wang and T. Lei. Statistical analysis of MR imaging and its applications in image modeling. *Proceedings of the IEEE International Conference on Image Processing and Neural Networks*, I, 1994, 866–870.