Chapter 3

Incremental Constructions along Space-Filling Curves

Summary. In this chapter we present an algorithm for incrementally constructing the Delaunay tessellation of a set of points. The algorithm combines randomness and locality by inserting the points in random rounds and within rounds in an order given by a space-filling curve.

We begin by generalizing *incremental constructions con BRIO*, i.e., incremental constructions with a *biased randomized insertion order*. In particular, we consider arbitrary configuration spaces and obtain a bound on the expected running time depending on the degree bound of the configuration space. Furthermore, we bound the expected running time for arbitrary sampling ratios. These bounds do not depend on the point distribution.

We then analyze our algorithm, i.e., incremental constructions along space-filling curves, for independent identically distributed points drawn uniformly from a bounded convex region in the plane. We prove that the algorithm runs in linear expected time in this case. Other point distributions will be analyzed in Chapter 4 where we also describe our implementation and discuss experimental results.

Preliminary versions of the results presented in Section 3.4 have appeared in [22, 23].

3.1 Introduction

When devising an insertion order for the incremental construction of the Delaunay tessellation of a point set there are two seemingly conflicting goals. Inserting points randomly from the data avoids creating many simplices during the construction which are not part of the final Delaunay tessellation. In contrast, inserting points nearby allows taking advantage of geometric locality and locality of reference.

The randomized incremental construction follows the first approach. It is asymptotically optimal but performs poorly with modern memory hierarchies when used for large data sets as observed by Amenta, Choi, and Rote [3]. They showed how randomness can be reduced without changing the asymptotic performance by a *biased randomized insertion order*: Points are randomly assigned to rounds of insertion of increasing sizes, and within a round the order of insertion can be chosen. A similar approach was proposed independently* by Zhou and Jones [153].

This allows to use locality within the rounds, thus combining a local insertion strategy with the randomized construction. In the algorithm presented here the points of a round are inserted in an order given by a space-filling curve, i.e., using the space-filling curve heuristic for the traveling salesperson problem. For point location *walking* is used, i.e., the tessellation is traversed locally to locate a point. For the analysis we assume a *straight line walk*. A new point is located from the point inserted in the previous step walking from simplex to simplex along a straight line.

Since the algorithm locates points by traversing the Delaunay tessellation, it does not require an additional data structure for point location. Nonetheless, such a data structure might be wanted, e.g., for locating points in the final Delaunay tessellation or as a fallback to speed up point location. We will therefore also discuss variants of the algorithm with an additional data structure for point location.

We choose a space-filling curve order because it creates a short tour through the points of the rounds, is fast to compute, and gives a good point distribution within a round. Furthermore it combines locality of reference with geometric locality by linearizing space, adapts well to irregularities of the point distribution, and is applicable in higher dimensions.

Our main goal is to present a theoretical argument for the strength of the combination of local and random insertion. We do this by giving an average case analysis of the algorithm. It runs in linear expected time for points distributed independently and uniformly in a bounded convex region. This proves that a biased randomized insertion order together with a local insertion scheme runs in linear expected time on uniform points, resolving an open problem posed by Amenta, Choi, and Rote [3]. This result complements the good practical performance of biased randomized insertion orders [3, 79, 95, 153].

The main technical contribution in the analysis of point location by walking is the explicit analysis near the boundary of a bounded convex region in the plane. For algorithms based on incremental construction, points near the boundary seem difficult to handle because long and thin triangles slow down the point location. Figure 3.1 shows a typical case of this: Near the boundary, triangles with a large circumcircle are likely to occur in the tessellation because a large part of the circumcircle may lie outside the region with points. The central part of the analysis in this chapter is to prove that the boundary case does not change the overall linearity.

Related Algorithms. Incremental constructions of Delaunay tessellations in two dimensions running in experiments in linear time on points generated according to a uniform distribution are given by Ohya, Iri and Murota [113, 114] using quaternary tree bucketing and by Su and Drysdale [139] using spiral

*S. Zhou, personal communication.

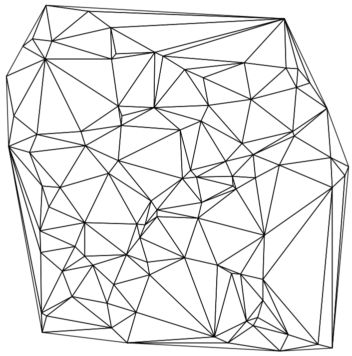


Figure 3.1: Delaunay tessellation of uniform points in a square.

search. In both cases, the analysis does not treat the irregularities near the boundary [114, 138]. Ohya et al. avoid the boundary case by considering points from a Poisson point process in the analysis. We will analyze a variant of the algorithm of Su and Drysdale in Section 4.6.

Linear expected time algorithms for constructing the Delaunay tessellation of independent uniformly distributed points from a bounded convex region in the plane based on other construction paradigms are known. Bentley, Weide and Yao [12] give an algorithm for constructing the Voronoi diagram which constructs each cell of the diagram separately by a spiral search. They prove that their algorithm runs in linear expected time in a bounded convex region in two dimensions. From the Voronoi diagram of linear complexity the Delaunay tessellation can be computed in linear time. Katajainen and Koppinen [84] use a divide-and-conquer approach for constructing the Delaunay tessellation for points in a square.

Dwyer [57] uses an algorithm for constructing the Delaunay tessellation based on incremental search. The algorithm constructs the Delaunay tessellation of independent uniformly distributed points drawn from a ball in arbitrary dimension in linear expected time. Incremental search starts with a known facet of the Delaunay tessellation, e.g., by computing the convex hull and taking a convex hull facet. Then the Delaunay tessellation is constructed by incrementally finding simplices neighboring known facets. This is achieved by searching the space near a known facet using a grid data structure. Constructing the Delaunay triangulation in the plane using incremental search has been analyzed by Maus [100]. Constructing the Voronoi diagram using incremental search has been considered by Tanemura, Ogawa and Ogita [145].

To obtain a linear expected running time, all of these algorithms assume linear time bucketing as discussed in the previous chapter, i.e., n points can be assigned to cubes in a cubical subdivision of size n in linear time. We will also use this assumption.

Independently to the research presented here, Liu and Snoeyink [95] (cf. [79]) have considered incrementally constructing the Delaunay tessellation along space-filling curves from a practical point of view. Instead of random sampling, they sample points to rounds based on the bit pattern of the least significant bits of the point coordinates. This reduces the bit complexity of the computation. They compare their algorithm with four current implementations for constructing the Delaunay tessellation in three dimensions. In the experimental comparison their implementation is the fastest on points representing the coordinates of atoms in proteins and on uniformly distributed points. Amenta, Choi, and Rote [3] use in their experiments oct-trees and *kd*-trees for computing orders within the rounds of incremental constructions con BRIO. Traversing an oct-tree yields a space-filling curve order. They report a good running time on surface data. Zhou and Jones [153] consider *hierarchical space-filling curve orders* and report a good running time on terrain data.

General Position. In the following chapters in which we discuss the construction of Delaunay tessellations we will assume that the points are in general position, i.e., no $d+1$ points lie in a $(d-1)$-flat and no $d+2$ points lie on a $(d-1)$-sphere. For instance, in two dimensions no 3 points lie on a line and no 4 points on a circle. Under this assumption the Delaunay tessellation is a simplicial complex for $n > d$. This assumption is not needed for the probabilistic analysis since they are implied by the probabilistic setup, i.e., the points are in general position with probability 1. We will assume further that the construction starts with a large bounding simplex containing all points. This simplifies the analysis in the case that a new point lies outside the convex hull of the points inserted so far. The techniques used in the algorithm generalize to the case where these assumptions do not hold (see Shewchuk's lecture notes [132]).

3.2 Preliminaries

We begin with reviewing *Voronoi diagrams* and *Delaunay tessellations*. The two main ingredients of the algorithm presented in this chapter are partially randomized incremental constructions, i.e., *incremental constructions con BRIO* [3], and space-filling curve orders. We will focus on incremental constructions without an additional data structure for point location, i.e., on *point location by walking*.

Space-filling curve orders have been discussed in the previous chapter. In the following we give background on the other concepts.

3.2.1 Voronoi Diagrams and Delaunay Tessellations

In Figure 1.1(b–c) we saw an example for a Voronoi diagram and the corresponding Delaunay tessellation in the plane. For a point set P the Voronoi diagram partitions the space into regions with common closest neighbor in the point set P, while the Delaunay tessellation links points in P if their Voronoi regions have a common face. The Voronoi diagram and the Delaunay tessellation depend on the underlying distance measure. In this and the following chapter we will use the Euclidean distance $\|x-y\|$.

The Voronoi diagram was first described informally by Descartes in 1644 in his Principia Philosophiae [44, 45]. It was for the first time formally introduced by Dirichlet in 1850 in two and three dimensions [54] and by Voronoi in 1908 in higher dimensions [148]. In 1934, Delaunay obtained the Delaunay tessellation by the *empty circle property* [43] (see Lemma 3.1) while Voronoi had considered it in his study of Voronoi diagrams. The Voronoi diagram has been rediscovered in other sciences [146, 152] and is known under various names. See [8, 59, 115] for further background on Voronoi diagrams and Delaunay tessellations.

Voronoi Diagram. Let P be a finite set of points in $\mathbb{R}^d$. The *Voronoi cell* of a point $p \in P$ is

$$V_p = \left\{ x \in \mathbb{R}^d \,\middle|\, (\forall q \in P)\ \|p - x\| \leq \|q - x\| \right\}.$$

Adjacent Voronoi cells share lower-dimensional parts, e.g., in two dimensions edges and vertices. We refer to the Voronoi cells and their lower dimensional parts as *Voronoi faces.* The collection of all Voronoi faces forms the Voronoi diagram.

Delaunay Tessellation. In the plane, the Delaunay tessellation has an edge for every pair of Voronoi cells which share a Voronoi edge. If the points are in general position, the faces enclosed by the Delaunay edges are triangles. Each Delaunay triangle corresponds to a Voronoi vertex. The vertices of the Delaunay tessellation, i.e., the original point set, correspond to Voronoi cells. In general in $\mathbb{R}^d$, for every k-dimensional Voronoi face there is a corresponding $(d-k)$-dimensional *Delaunay face.* For points in general position the Delaunay faces are simplices. In the plane the two-dimensional faces are triangles. We will refer to the Delaunay tessellation in the plane as Delaunay triangulation.

In the plane, the 1-skeletons of a Delaunay triangulation, i.e., the graph formed by its vertices and edges, is a planar graph. Therefore, by Euler's formula its average degree is less than 6. In particular, the complexity of a Delaunay triangulation and of a Voronoi diagram in the plane is linear in the number of points. In higher dimensions the complexity of the Delaunay tessellation and the Voronoi diagram is bounded by $O(n^{\lceil \frac{d}{2} \rceil})$ and this bound is tight in the worst case [86, 127].

A subset of points $S \subset P$ of a point set P forms a simplex in the Delaunay tessellation if and only if it also defines a Voronoi face. This is equivalent to the existance of a point which has exactly the points in S as nearest neighbors. This is the case if there is a d-ball around this point with exactly the points of S on the corresponding sphere and no further points of P in the ball. Under general position assumption, if there is a ball with empty interior and the points in S on its sphere, then there is also an empty ball with exactly the points in S on its sphere. Thus, we have the following lemma.

Lemma 3.1 (empty circumsphere property)**.** *Let P be a finite point set in $\mathbb{R}^d$ in general position. A subset $S \subset P$ of points forms a simplex in the Delaunay tessellation of P if and only if there is a d-ball through the points in S with no further points of P inside.*

3.2.2 Randomized Incremental Construction

Incremental Construction. The basic concept of incremental construction is simple to state: Insert the points into the Delaunay tessellation one by one, updating the data structure after each insertion step. The time needed to insert a point consists of the time needed for *locating the point* in the current tessellation and the time for *updating* the tessellation. The running time of the randomized incremental construction of the Delaunay tessellation can be analyzed in the framework of *configuration spaces* (see Section 3.2.3).

Update. In the update step the algorithm updates the Delaunay tessellation starting at a simplex conflicting with the new point, i.e., at a simplex with the new point in its circumsphere. There are two common techniques for updating, *flipping* and *Bowyer-Watson updates*.

Edge flips in the construction of Delaunay triangulations in two dimensions were introduced by Lawson [89]. The triangulation is updated by first splitting the triangle which contains the new point into three triangles. The resulting triangulation is not necessarily the Delaunay triangulation of the point set because conflicts between triangles and vertices of the triangulation might have been created. These can be resolved by flipping the diagonals of quadrilaterals, i.e., of two adjacent triangles, if one of the triangles (and therefore also the other) is in conflict with the opposite vertex of the quadrilateral. Starting at the triangles which contain the new point, quadrilaterals at newly created triangles are checked (and possibly updated) until there are locally no more conflicts. Then also globally there are no more conflicts and the resulting triangulation is the Delaunay triangulation. Flipping can be generalized to higher dimensions. See [59] for more details on flipping.

For the Bowyer-Watson update [20, 150] the tessellation is searched locally for conflicting simplices. These are all removed and the resulting void region is tessellated by the simplices which are obtained by connecting the new vertex to the lower-dimensional simplices on the boundary of the region. The resulting tessellation is the Delaunay tessellation. See Shewchuk's lecture notes [132] for more details on Bowyer-Watson updates.

For both types of updates, the time needed for updating is asymptotically linear in the number of simplices removed and added. Thus, in total it is linear in the number of simplices occurring during the construction.

Point Location by Walking. In the point location step the algorithm determines a simplex from which to start the update, i.e., a simplex which is in conflict with the new point. A point can be located directly in the data structure of the tessellation by *walking* or in an additional data structure.

Point location by walking finds a new point by locally traversing the data structure of the Delaunay tessellation [73, 89]. For an efficient implementation of walking, the data structure must provide access to neighboring simplices in constant time. An example of such a data structure is the cell-tuple structure [21]. In this chapter we will consider straight line walks: Starting from a vertex of the tessellation, the new point is located by traversing the simplices intersected by a straight line between the vertex and the new point. A different walking scheme, *Lawson's oriented walk*, will be considered in Section 4.6. Further walking schemes are discussed by Devillers, Pion, and Teillaud [47]. The time needed for walking is linear in the number of simplices traversed.

In the worst case, the number of simplices traversed for locating a point can be linear in the complexity of the tessellation. However, for independent uniformly distributed points in a bounded convex region in the plane the expected number is smaller.

Lemma 3.2 (Devroye, Mücke and Zhu [50]). *The expected number of intersections between a Delaunay triangulation of points distributed independently and uniformly in a bounded convex region $C \subset \mathbb{R}^2$ of area 1 and a fixed line segment L that is at least a distance $c_0\sqrt{\log n/n}$ from the boundary of C is bounded by*

$$c_1|L|\sqrt{n} + c_2,$$

where c_0 is an absolute constant, and c_1 and c_2 depend only on geometric properties of C.

Devroye, Lemaire and Moreau [49] present a proof of this lemma for the case of the unit square. Katajainen and Koppinen [84] give a similar bound in the analysis of their divide-and-conquer algorithm for constructing the Delaunay triangulation. Bose and Devroye [19] bound the number of intersections between random Voronoi diagrams and line segments. As a corresponding result in three dimensions Mücke, Saias and Zhu [106] give a $O\left((1+|L|)n^{1/3}\frac{\log n}{\log\log n}\right)$ bound on the number of intersections in $\mathbb{R}^3$ for line segments with distance at least $c(\log n/n)^{1/3}$ to the boundary for a constant c depending on the region C.

We will need to extend Lemma 3.2 later (Lemma 3.25) and will bound the constants c_0, c_1, and c_2 there. The condition of L being a fixed line segemnt can be relaxed to L being a random line segment independent of the points in the triangulation by using conditional probabilities.

Lemma 3.2 yields that if points are inserted by starting at a random previously inserted point (e.g., since the points are independent this might be the point inserted last) the expected cost for point location is $O(\sqrt{n})$. Near the boundary the expected cost is higher but this cost can be amortized since there are only few points near the boundary. The expected total running time is $O(n^{3/2})$. This can be reduced to $O(n^{4/3})$ by starting the point location not at a random point but at the nearest neighbor of the new point out of a random sub-sample of size $\Theta(n^{1/3})$.

To prove that the incremental construction along space-filling curves has linear expected running time we need to prove that the expected cost of walking along a space-filling curve tour takes expected constant time per point. For the two-dimensional case we will use Lemma 3.2 but with a refined analysis near the boundary.

Point Location with Data Structure. Various data structures for point location together with a random insertion order lead to algorithms for constructing the Delaunay tessellation which are asymptotically optimal for worst-case point sets taking the average over the insertion orders. Here we discuss the conflict graph, the history graph, and the Delaunay hierarchy.

The *conflict graph* [37] is a bipartite graph with vertices for the d-simplices of the tessellation and for points not yet inserted. The edges represent conflicts between the points and simplices. For point location in a Delaunay tessellation fewer edges suffice: the *simplified conflict graph* [37, 132] only maintains an edge between a point not yet inserted and the simplex containing it. At the beginning of the algorithm the graph contains an edge between each point and the bounding simplex. During the construction the simplified conflict graph allows to find the simplex containing a new point in constant time. Points lying in simplices that are destroyed must be redistributed to the simplices newly created.

An alternative to the conflict graph is the *history graph*, also called *history dag* (directed acyclic graph) or *influence graph* [18, 76]. Its main advantage over the conflict graph is that it needs to access a point not until it is inserted while the conflict graph needs to access all points from the beginning of the construction on (until it has been inserted). The history is a directed acyclic graph. It has a vertex for every simplex that was created during the construction. Initially, it contains only the bounding simplex. If a simplex is destroyed during the construction, an edge from this simplex to the simplices replacing it is added to the history. A point is located using the history by starting with a bounding simplex and by traversing the sequence of simplices containing the point.

The sequence of simplices traversed for a point in the history is the same as for the simplified conflict graph although this sequence is traversed at different times in the two versions of the algorithm. Thus, the running time of point location with the (simplified) conflict graph and with the history are asymptotically equal. When a simplex is added or removed from the tessellation then an edge is added or removed for every point in conflict with the simplex. In the history graph a node will be traversed whenever a point is located that is in conflict with the simplex corresponding to the node. Both the (simplified) conflict graph and the history can be implemented such that the total running time for point location is asymptotically equal to the sum of the number of conflicting points for all simplices [37]. The running time can be analyzed in the general framework of configuration spaces (see Section 3.2.3).

A further expected worst-case optimal data structure for constructing the Delaunay tessellation using randomized incremental construction is the *Delaunay hierarchy* [46]. It stores the Delaunay tessellations for a nested sequence of random samples of the current point set. For a point present in the current tessellation it stores a link to a simplex incident to the point in all tessellations of sub-samples in which it is present. To insert a new point, it is first determined to which samples it is added. If the random samples are denoted by

$$\mathcal{S} = \mathcal{S}_0 \supseteq \mathcal{S}_1 \supseteq \mathcal{S}_2 \supseteq \dots$$

then the samples in which a new point p is included is determined by Bernoulli sampling

$$\mathrm{P}\left[p \in \mathcal{S}_{i+1} \,|\, p \in \mathcal{S}_i\right] = \frac{1}{\alpha}$$

for $i \geq 0$ and a sampling parameter $\alpha > 1$. A new point is located by successively locating its nearest neighbor in the Delaunay tessellations of the samples starting with the smallest sample. The expected running time for constructing the Delaunay tessellation using the Delaunay hierarchy is worst-case optimal [46].

3.2.3 Configuration Spaces

Randomized incremental constructions can often be analyzed using the framework of configuration spaces. A configuration space (cf. [35]) consists of

1. a finite set of geometric objects X,
2. a set of configurations $\Pi(X)$,
3. a mapping assigning to every $S \subset X$ a set of configurations $\mathcal{T}(S) \subset \Pi(X)$,
4. mappings $\mathcal{D}, \mathcal{K}$ assigning to every configuration Δ sets of objects in X: $\mathcal{D}(\Delta) \subset X$ are called *triggers* of Δ, $\mathcal{K}(\Delta) \subset X$ are called *stoppers* of Δ. The number of triggers of Δ is called *degree* of Δ, the number of stoppers is called *conflict size*.

For the bounds presented in this section, a configuration space must fulfill the following conditions:

- for every configuration Δ and every $S \subset X$ it holds that $\Delta \in \mathcal{T}(S)$ if and only if $\mathcal{D}(\Delta) \subset S$ and $\mathcal{K}(\Delta) \cap S = \emptyset$,
- $\max_{\Delta \in \Pi(X)}\{|\mathcal{D}(\Delta)|\}$ is bounded by a constant δ_0, called the *degree bound* of the configuration space.

In our case we have the following configuration space:

1. the set X of objects is a finite set of points $P \subset \mathbb{R}^d$,
2. the set of configurations $\Pi(X)$ is the set of all simplices defined by P,
3. the set of configurations $\mathcal{T}(S)$ assigned to a subset $S \subset P$ is its Delaunay tessellation $\mathcal{DT}(S)$,
4. the triggers of a simplex are its vertices and the stoppers are all points in the interior of its circumsphere.

The two conditions above are fulfilled, in particular all triggers have size $d + 1$ by the general position assumption.

The expected update and point location cost (using the history) for the randomized incremental construction of the Delaunay tessellation can be bounded using the framework of configuration spaces. Recall that the update cost is asymptotically bounded by the number of simplices occurring during the construction, called *total structural change* and denoted here by $C_u(n)$. Here n denotes the size of X, i.e., the number of points in the case of the Delaunay tessellation. The point location cost is bounded by summing up the number of conflicting points for all simplices occurring during the construction, called *total conflict size* and denoted here by $C_\ell(n)$.

The following bounds on C_u and C_ℓ hold (using the notions for configuration spaces) [37]:

$$\mathrm{E}\left[C_u(n)\right] \leq \sum_{i=1}^{n} \frac{\delta_0}{i} \mathrm{E}\left[|\mathcal{T}(X^i)|\right] \quad \text{and} \tag{3.1}$$

$$\mathrm{E}\left[C_\ell(n)\right] \leq \sum_{i=1}^{n} \frac{\delta_0^2(n-i)}{i^2} \mathrm{E}\left[|\mathcal{T}(X^i)|\right], \tag{3.2}$$

where X^i denotes a random sample of X of size i for $1 \leq i \leq n$.

Thus, if we assume the dimension d to be constant and use the worst-case bound of $i^{\lceil \frac{d}{2} \rceil}$ on $|\mathcal{T}(X^i)|$ this yields

$$\mathrm{E}\left[C_u(n)\right] \in O(n^{\lceil \frac{d}{2} \rceil}), \tag{3.3}$$

$$\mathrm{E}\left[C_\ell(n)\right] \in O(n \log n + n^{\lceil \frac{d}{2} \rceil}). \tag{3.4}$$

We will consider point distributions for which the expected size of the Delaunay tessellation is linear in the number of points in the tessellation. For these distributions the inequalities (3.1) and (3.2) yield

$$\mathrm{E}\left[C_u(n)\right] \in O(n), \tag{3.5}$$

$$\mathrm{E}\left[C_\ell(n)\right] \in O(n \log n). \tag{3.6}$$

We are interested in algorithms running in linear expected time on the distributions considered. By bound (3.5) the expected update cost is linear. For the expected point location cost we so far only have the super-linear bound (3.6). We will therefore focus on the point location cost.

Let k_s be the number of simplices in $\Pi(P)$ with exactly s stoppers and let p_s be the probability that a given simplex with s stoppers appears during the construction. Then $\mathrm{E}\left[C_u(n)\right]$ and $\mathrm{E}\left[C_\ell(n)\right]$ can be rewritten as

$$\mathrm{E}\left[C_u(n)\right] = \sum_{s=0}^{n} k_s p_s \quad \text{and}$$

$$\mathrm{E}\left[C_\ell(n)\right] = \sum_{s=0}^{n} s k_s p_s.$$

The bounds (3.3 – 3.6) can also be derived directly using these equations in combination with the following bound by Clarkson and Shor on the number K_s of configurations in $\Pi(X)$ with at most s stoppers:

Theorem 3.3 (Clarkson, Shor [37]).

$$K_s \in O(s^{\delta_0} E[|\mathcal{T}(R)|]),$$

where R is a random sample of X of size $\lfloor |X|/s \rfloor$.

In the following we derive the bounds (3.3 − 3.6) from Theorem 3.3. A similar derivation was given in the original analysis of incremental constructions con BRIO [3]. Our analysis of BRIOs in Section 3.3.1 will not require such a derivation since we give bounds for the incremental constructions con BRIO relative to randomized incremental constructions. For completeness, we give the derivation here.

$$\begin{aligned} \mathrm{E}\,[C_u(n)] &= \sum_{s=0}^{n} k_s p_s \\ &= k_0 p_0 + \sum_{s=1}^{n} (K_s - K_{s-1}) p_s \\ &= k_0 + \sum_{s=1}^{n-1} K_s (p_s - p_{s+1}) + K_n p_n \\ &\leq k_0 + c \sum_{s=1}^{n} s^{d+1} \mathrm{E}\left[|\mathcal{T}(X^{\frac{n}{s}})|\right] (p_s - p_{s+1}) \end{aligned}$$

for suitable n and c, $p_{n+1} := 0$, and $X^{\frac{n}{s}}$ a random sample of $\lfloor \frac{n}{s} \rfloor$ points.

Consider the event that a configuration with δ triggers and s stoppers appears in a randomized incremental construction. The probability $p_R(\delta, s)$ of this event is the probability for choosing δ triggers from $\delta + s$ triggers and stoppers which is

$$p_R(\delta, s) = \frac{1}{\binom{\delta+s}{\delta}}. \tag{3.7}$$

Now,

$$\begin{aligned} p_s - p_{s+1} &= \frac{1}{\binom{s+d+1}{d+1}} - \frac{1}{\binom{s+d+2}{d+1}} = \frac{(d+1)!s!(d+1)}{(s+d+2)!} \\ &< \frac{1}{\binom{s+d+2}{s}} < \frac{1}{s^{d+2}}. \end{aligned}$$

Thus,

$$\mathrm{E}\,[C_u(n)] \leq k_0 + c \sum_{s=1}^{n-1} \mathrm{E}\left[|\mathcal{T}(X^{\frac{n}{s}})|/s\right].$$

By the same derivation we get

$$\mathrm{E}\,[C_\ell(n)] \leq k_0 + c \sum_{s=1}^{n-1} s^{d+1} \mathrm{E}\left[|\mathcal{T}(X^{\frac{n}{s}})|\right] (s\, p_s - (s+1) p_{s+1}).$$

Since

$$s\, p_s - (s+1) p_{s+1} < s(p_s - p_{s+1}) < \frac{1}{s^{d+1}}$$

Algorithm 3: Incremental Construction con BRIO

Input: Point set in $\mathbb{R}^3$
Output: Delaunay tessellation of the point set

1 Compute BRIO:
1.1 Sample points to rounds (using coin flips),
1.2 Order points in a round (any order can be used).
2 Incrementally construct Delaunay tessellation using order from step 1: In each step do
2.1 Locate new point,
2.2 Update Delaunay tessellation.

we get

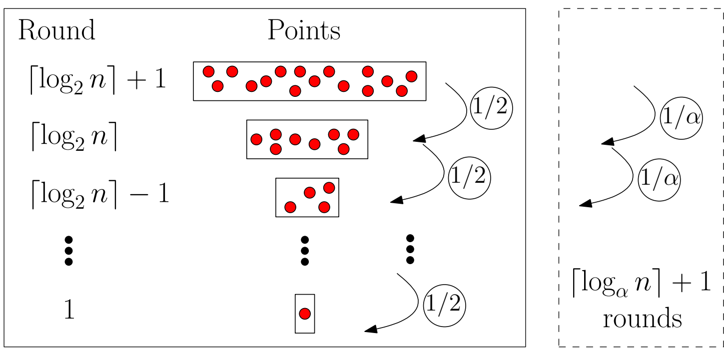
$$\mathrm{E}\left[C_\ell(n)\right] \;\leq\; k_0 + c\sum_{s=1}^{n-1} \mathrm{E}\left[\left|\mathcal{T}(X^{\frac{n}{s}})\right|\right].$$

From this, the bounds (3.3 − 3.6) above follow.

3.2.4 Incremental Constructions con BRIO

Biased randomized insertion orders (BRIOs) where proposed by Amenta, Choi, and Rote [3] for incrementally constructing the Delaunay tessellation of points in three dimensions, see Algorithm 3. Sufficient randomness can be introduced to the insertion order by assigning the points independently at random to rounds, where the number of points in one round is approximately the same as the number of points in all previous rounds.

The sampling is illustrated in Figure 3.2(a): A point is independently assigned to the last round with the probability of 1/2. Each of the remaining points is assigned to the next to last round with the probability of 1/2, and so on [3]. After a logarithmic number of rounds an expected constant number of points remain, and we can therefore stop the sampling and assign the remaining points to the first round. If $p \in R_{\leq i}$ denotes that the point p is inserted in round i or before ($i \geq 1$) then the assignment can be described in terms of



(a) (b)

Figure 3.2: Assigning points to rounds.

probabilities as

$$\mathrm{P}\left[p \in R_{\leq i} \,|\, p \in R_{\leq i+1}\right] = \frac{1}{2}$$

for $1 \leq i < \lceil \log_2 n \rceil + 1$ and

$$\mathrm{P}\left[p \in R_{\leq \lceil \log_2 n \rceil + 1}\right] = 1.$$

This yields the following distribution of points to rounds if n is a power of 2: A point is assigned to the first round with probability $1/n$, and to the kth round with probability $2^{k-2}/n$ for $2 \leq k \leq \lceil \log_2 n \rceil + 1$. Points are inserted round by round. Within one round points can be inserted in an arbitrary order without changing the asymptotic running time for insertion.

Biased randomized insertion orders were originally introduced to avoid non-local and completely random memory access patterns. We make use of the fact that they do not change the update cost, which is linear in our case. Therefore we can focus on the time for point location. For us the main objective is to balance randomness and geometric locality to speed up point location.

For many input point sets the expected complexity of the Delaunay tessellation of a random sample of the point set is linear in the size of the sample. Amenta, Choi, and Rote call this the *realistic case*. For incremental constructions con BRIO they prove the following theorem:

Theorem 3.4 (Amenta, Choi, and Rote [3]). *With incremental construction con BRIO the expected total number of tetrahedra that are created during the construction of the Delaunay tessellation of n points in three dimensions is $O(n^2)$ in the worst case and $O(n)$ in the realistic case. The expected total conflict size (and hence the expected running time with a history) is $O(n^2)$ in the worst case and $O(n \log n)$ in the realistic case.*

3.3 Incremental Constructions con BRIO Revisited

Incremental constructions con BRIO were formulated and analyzed for constructing the Delaunay tessellation in three dimensions using a history or conflict graph [3]. For our purposes we need to extend the algorithm in several ways.

First, we want to allow an arbitrary sampling ratios $1/\alpha$ with $\alpha > 1$, i.e., with a biased coin instead of fair coin (see Figure 3.2(b)). In the following we will always describe the sampling ratio $1/\alpha$ by the *sampling parameter* α. We impose a bound of $k_\alpha := \lceil \log_\alpha n \rceil + 1$ on the maximum number of rounds. Points are assigned independently at random to rounds such that for every point p

$$\mathrm{P}\left[p \in R_{\leq i} \,|\, p \in R_{\leq i+1}\right] = \frac{1}{\alpha}$$

for $1 \leq i < k_\alpha$ and

$$\mathrm{P}\left[p \in R_{\leq k_\alpha}\right] = 1.$$

This is motivated by the observation that in our experiments the sampling parameter 2 (as in the original construction) is not optimal. The speed-up arises because the cost for point location by walking depends on α.

Algorithm 4: Incremental Construction con BRIO revisited

Input: Objects from a configuration space

Output: Set of configurations corresponding to the set of objects

1 Compute BRIO:

1.1 Sample objects to rounds (using coin flips with sampling parameter $\alpha > 1$),

1.2 Order objects in a round (any order can be used).

2 Incrementally construct set of configurations using the order from step 1.

Second, we want to use the construction in arbitrary dimension. For this we extend the analysis of incremental constructions con BRIO to arbitrary configuration spaces, see Algorithm 4. Furthermore, we simplify the analysis by giving bounds on the expected costs relative to the expected costs of a randomized incremental construction instead of computing them explicitly. A bound for the expected running time depending on the sampling parameter and the configuration space is given in Proposition 3.7. A refined bound is given in Theorem 3.8.

Third, we want to estimate the time needed to construct and use the history of a single round. This is motivated by the following question. Assume we can locate the points of the current round efficiently in the Delaunay tessellation of the points of the previous rounds. Can they then be efficiently located with the history in the current tessellation at the time they are inserted? We will show in Theorem 3.9 that this is the case. Our main algorithm however will not require the history or any other point location data structure.

Notation. We will denote the expectation of a construction with sampling parameter α by $\mathrm{E}_{BRIO(\alpha)}[\cdot]$ if we want to contrast it to the expectation of a randomized incremental construction, denoted by $\mathrm{E}_{RIC}[\cdot]$. We will also consider incremental constructions con BRIO without a bound on the maximum number of rounds, i.e., with $k_\alpha = \infty$. In this case we denote the expectation by $\mathrm{E}_{BRIO^*(\alpha)}[\cdot]$.

Consider the event that a configuration with δ triggers and s stoppers appears in the ith to last round using a BRIO without bound restriction, for $i \geq 1$ where $i = 1$ refers to the last round. This event can occur only if the following two conditions hold:

(1) all stoppers are in the ith to last round or later,

(2) all triggers are in the ith to last round or before, and not all triggers are before the ith round.

Therefore the probability for the event to occur can be bounded by the probability $p_i(\delta, s)$ for conditions (1) and (2) to hold. Condition (1) is necessary because a configuration cannot occur if any of its stoppers have been inserted previously. Condition (2) is necessary because a configuration can only occur in the round in which its last trigger is inserted. The probability for a point to be in the ith to last round or before is $\frac{1}{\alpha^i}$. Thus, the probability for condition

(1) is $\left(1-\frac{1}{\alpha^i}\right)^s$, for condition (2) is $\left(1-\frac{1}{\alpha^{i-1}}\right)^\delta - \left(1-\frac{1}{\alpha^i}\right)^\delta = \left(1-\frac{1}{\alpha^\delta}\right)\alpha^{\delta(1-i)}$, and therefore

$$p_i(\delta, s) = \left(1-\frac{1}{\alpha^i}\right)^s \left(1-\frac{1}{\alpha^\delta}\right)\alpha^{\delta(1-i)}. \tag{3.8}$$

The probability for the configuration to appear in the total construction using a BRIO without bound restriction is therefore bounded by

$$p_B(\delta, s) := \sum_{i=1}^{\infty} p_i(\delta, s). \tag{3.9}$$

3.3.1 Con BRIO Generalized and Simplified

We prove that if the degree bound δ_0 of the configuration space and the sampling parameter α are considered as constants then the expected total structural change $\mathrm{E}_{BRIO(\alpha)}[C_u(n)]$ and the expected total conflict cost $\mathrm{E}_{BRIO(\alpha)}[C_\ell(n)]$ are asymptotically bounded by the corresponding costs for a randomized incremental construction. The proof proceeds in two steps: In Lemma 3.5 we bound $\mathrm{E}_{BRIO^*(\alpha)}[C_u(n)]$ and $\mathrm{E}_{BRIO^*(\alpha)}[C_\ell(n)]$, i.e., the costs without a bound on the number of rounds. In Lemma 3.6 we bound the error terms due to a bound on the number of rounds. In the proof of both lemmas we will consider for a configuration Δ the following events:

- $T_i :=$ {All triggers of Δ appear in the ith to last round or before.},
- $S_i =$ {The first stopper of Δ appears in the ith to last round.},

where $i \in \mathbb{N}$, and $i = 1$ refers to the last round. For instance, S_2 is the event that all stoppers of Δ are inserted in the last two rounds and at least one stopper is inserted in the second to last round. The disjoint union $\sum_{i=1}^{\infty} S_i$ has probability 1. If S_i holds for an $i \in \mathbb{N}$ then Δ can only appear during the construction if also T_i holds.

Lemma 3.5. *For a sampling parameter $\alpha > 1$ and a configuration space with degree bound δ_0 the following inequalities hold:*

$$\begin{aligned} E_{BRIO^*(\alpha)}[C_u(n)] &\le \alpha^{\delta_0} E_{RIC}[C_u(n)], \\ E_{BRIO^*(\alpha)}[C_\ell(n)] &\le \alpha^{\delta_0} E_{RIC}[C_\ell(n)]. \end{aligned}$$

Proof. It suffices to prove for all pairs of numbers of triggers and stoppers (δ, s)

$$p_B(\delta, s) \le \alpha^\delta p_R(\delta, s).$$

Then we have that $p_B(\delta, s) \le \alpha^{\delta_0} p_R(\delta, s)$ since $\alpha > 1$ and $\delta \le \delta_0$. Thus, the bounds hold by linearity of expectation.

For $s = 0$ we have

$$p_B(\delta, s) = 1 = p_R(\delta, s) \le \alpha^\delta p_R(\delta, s).$$

Assume $\Delta_{\delta,s}$ is a configuration with δ triggers and $s > 0$ stoppers. Then,

$$p_B(\delta, s) \;\leq\; \mathrm{P}\left[\sum_{i=1}^{\infty}(S_i \cap T_i)\right] \;=\; \sum_{i=1}^{\infty}\mathrm{P}\left[S_i \cap T_i\right] \;=\; \sum_{i=1}^{\infty}\mathrm{P}\left[S_i\right]\mathrm{P}\left[T_i\right]$$

where the last equality holds by the independence of S_i and T_i for $i \in \mathbb{N}$. Now,

$$\mathrm{P}\left[T_{i+1}\right] \;=\; \mathrm{P}\left[T_{i+1} \cap T_i\right] \;=\; \mathrm{P}\left[T_{i+1}\,|\,T_i\right]\mathrm{P}\left[T_i\right] \;=\; \frac{1}{\alpha^{\delta}}\mathrm{P}\left[T_i\right]$$

by the sampling condition. Therefore,

$$p_B(\delta, s) \;\leq\; \sum_{i=1}^{\infty}\alpha^{\delta}\mathrm{P}\left[S_i\right]\mathrm{P}\left[T_{i+1}\right] \;=\; \alpha^{\delta}\mathrm{P}\left[\sum_{i=1}^{\infty}(S_i \cap T_{i+1})\right],$$

i.e., $p_B(\delta, s)$ is bounded by α^{δ} times the probability that all triggers are inserted in rounds strictly after the stoppers. The probability for all triggers to be inserted in rounds strictly after the stoppers does not depend on the order within rounds. Consider an incremental construction con BRIO with random order within rounds. This gives a randomized incremental construction. The event that all triggers appear in rounds before all stoppers is included in the event that all triggers appear before all stoppers. Thus, we get

$$\mathrm{P}\left[\sum_{i=1}^{\infty}(S_i \cap T_{i+1})\right] \;\leq\; p_R(\delta, s) \tag{3.10}$$

and therefore $p_B(\delta, s) \;\leq\; \alpha^{\delta} p_R(\delta, s)$. □

In the last step in the proof above, i.e., inequality (3.10), we could alternatively argue directly that the probability is bounded by $1/\binom{s+\delta}{\delta}$ in the same way as for $p_R(\delta, s)$ (equation (3.7)). Note that equation 3.10 does not hold with equality, but we will see (Theorem 3.8) that the bound is asymptotically tight up to a factor $\frac{1}{\sqrt{\delta}}$.

Lemma 3.6. *For a sampling parameter $\alpha > 1$ and a configuration space with degree bound δ_0 the following inequalities hold:*

$$\begin{aligned} E_{BRIO(\alpha)}\left[C_u(n)\right] &< E_{BRIO^*(\alpha)}\left[C_u(n)\right] + e, \\ E_{BRIO(\alpha)}\left[C_\ell(n)\right] &< E_{BRIO^*(\alpha)}\left[C_\ell(n)\right] + en, \end{aligned}$$

where e denotes Euler's constant.

Proof. Let Δ be a configuration with δ triggers and s stoppers. As in the previous proof we can bound the probability $p_B(\delta, s)$ of Δ occurring during the construction by

$$p(\delta, s) \;\leq\; \sum_{i=1}^{k_\alpha}\mathrm{P}\left[S_i\right]\mathrm{P}\left[T_i\right]$$

where the sum now only runs to $k_\alpha = \lceil \log n / \log \alpha \rceil + 1$. All summands except the last can be bounded as before. For the last summand we get

$$\mathrm{P}\left[S_{k_\alpha}\right]\mathrm{P}\left[T_{k_\alpha}\right] \;\le\; \mathrm{P}\left[T_{k_\alpha}\right] \;\le\; \frac{1}{\alpha^{\lceil \log_a n \rceil + 1}} \;\le\; \frac{1}{n^\delta} \;=:\; p'(\delta, s).$$

For a degree $\delta \le \delta_0$ there are at most $\binom{n}{\delta} \le \frac{n^\delta}{\delta!}$ configurations. A configuration appearing during the construction contributes 1 to the total structural change and at most n to the total conflict size. Thus, the expected additional structural change induced by the event $S_{k_\alpha} \cap T_{k_\alpha}$ is bounded by

$$\sum_{\delta=0}^{\delta_0} \binom{n}{\delta} p'(\delta, s) \;<\; \sum_{\delta=0}^{\infty} \frac{1}{\delta!} \;=\; e$$

and the expected additional conflict size is bounded by

$$\sum_{\delta=0}^{\delta_0} n \binom{n}{\delta} p'(\delta, s) \;<\; en.$$

□

Combining Lemmas 3.5 and 3.6 gives the following proposition.

Proposition 3.7. *For a sampling parameter $\alpha > 1$ and a configuration space with degree bound δ_0 the following inequalities hold:*

$$\begin{aligned} E_{BRIO}\left[C_u(n)\right] &< \alpha^{\delta_0} E_{RIC}\left[C_u(n)\right] + e, \\ E_{BRIO}\left[C_\ell(n)\right] &< \alpha^{\delta_0} E_{RIC}\left[C_\ell(n)\right] + en, \end{aligned}$$

where e denotes Euler's constant.

3.3.2 Refined Analysis and Cost of Building the History

In the following we give a refined bound for the probability $p_i(\delta, s)$ that a configuration with δ triggers and s stoppers appears in the ith to last round. As before, we bound this probability and the worst case probability $p_B(\delta, s)$ of the configuration to appear in the construction con BRIO without bound on the number of rounds relative to the probability $p_R(\delta, s)$ that the configuration appears in the randomized incremental construction. The motivation for this is twofold. First, this allows to tighten the bounds obtained above. Second, this gives an estimate for the cost of building the history of a single round.

Consider the following point location scheme. At the beginning of a round all points of the round are located in the Delaunay tessellation of the points of the previous rounds – how this is achieved will be discussed later. Then within the round the points are located starting from this location using the history of this round. By doing this in every round, we build the history of the complete construction. We prove that the expected cost of building the history of the last round (including the cost of using the history for point location) is asymptotically bounded by the expected update cost.

The main results of the refined analysis are the following two theorems.

Theorem 3.8. *For a sampling parameter $\alpha > 1 + \frac{1}{\delta}$ and a configuration space that has only configurations of degree $\delta > 0$ the following inequalities hold:*

$$\begin{aligned} E_{BRIO^*}\left[C_u(n)\right] &< \frac{\alpha^\delta}{\sqrt{\delta}}\, c_{\alpha,\delta}\, E_{RIC}\left[C_u(n)\right] \\ E_{BRIO^*}\left[C_\ell(n)\right] &< \frac{\alpha^\delta}{\sqrt{\delta}}\, c_{\alpha,\delta}\, E_{RIC}\left[C_\ell(n)\right], \end{aligned}$$

with $c_{\alpha,\delta} := \frac{e^{13/12}}{\sqrt{2\pi}}\left(1 - \frac{1}{\alpha^\delta}\right)\frac{\alpha+1}{\alpha-1} \cdot \frac{\delta\alpha-1}{\delta(\alpha-1)-1}$ *and where e denotes Euler's constant.*

For incremental constructions con BRIO let $C_\ell^1(n)$ be the sum of the conflict sizes of all configurations appearing in the last round or existing at the beginning of the last round. Thus, $C_\ell^1(n)$ counts all configurations that have all stoppers in the last round weighted by the number of stoppers they have.

Theorem 3.9. *For a sampling parameter $\alpha > 1$ and a configuration space that has only configurations of degree $\delta > 0$ the following inequality holds:*

$$E_{BRIO^*}\left[C_\ell^1(n)\right] \quad < \quad \alpha^{\delta+1}\sqrt{\delta+1}\, c_{\alpha,\delta}\, E_{RIC}\left[C_u(n)\right],$$

with $c_{\alpha,\delta} := \frac{e^{1/12}}{\sqrt{2\pi}}\left(\frac{\alpha}{\alpha-1}\right)^{\frac{3}{2}}\left(\frac{\delta+1}{\delta}\right)^{\frac{1}{2}}$ *and where e denotes Euler's constant.*

Theorem 3.8 will follow from Lemma 3.14. Theorem 3.9 will follow from Lemma 3.11, Lemma 3.12 and Observation 3.13.

For large α, the bounds in Theorem 3.8 improve the bounds in Proposition 3.7 by a factor $1/\sqrt{\delta}$. We will see that the bounds in Theorem 3.8 and in Theorem 3.9 are best possible, in the sense that for sufficiently large α and up to a constant factor no better bounds can be proved based on a comparison of $p_B(\delta, s)$ and $p_R(\delta, s)$, and $s\, p_1(\delta, s)$ and $p_R(\delta, s)$, respectively. This will follow from Lemmas 3.11 and 3.12.

For Delaunay tessellations, Theorem 3.9 bounds the expected cost of building the history of the last round after locating the points of the last round in the Delaunay tessellation of the points of the previous round. By Theorem 3.9 the expected cost of building the history of the last round is asymptotically bounded by the update cost of a randomized incremental construction (up to a constant depending on α and δ) and therefore by the expected complexity of the tessellation. If the expected complexity of the Delaunay tessellation is linear in the number of points then by the linearity of expectation the expected total cost of building the history in this way, i.e., building the history in all rounds, is linear. Thus, Theorem 3.9 implies the following corollary.

Corollary 3.10. *Let $P \subset \mathbb{R}^d$ be a set of n points such that the expected complexity of the Delaunay tessellation of a random sample is linear in the size of the sample. In an incremental construction con BRIO of the Delaunay tessellation of P the expected cost of building the history and locating points in the history is in $O(n)$, if before each round the points of the round are located in the current Delaunay tessellation.*

Probability of Appearing in a Round. We start with bounding the probability that a configuration appears in a round relative to the probability that it appears in a randomized incremental construction. For the analysis, we consider the parameter s of the probabilities $p_R(\delta, s), p_B(\delta, s)$, and $p_i(\delta, s)$ for $i \geq 1$ as real number, and denote the generalized probabilities by $\bar{p}_R(\delta, s), \bar{p}_B(\delta, s)$, and $\bar{p}_i(\delta, s)$ (using formulas (3.7)-(3.9)).

Lemma 3.11. *For $\alpha > 1$, $\delta \geq 1$, and $i \geq 1$ holds*

$$\max_{s>0} \frac{\bar{p}_i(\delta, s)}{\bar{p}_R(\delta, s)} \quad = \quad \frac{\alpha^\delta}{\sqrt{\delta}}\, c_{\alpha,\delta,i}$$

where

$$c_{\alpha,\delta,i} \quad \geq \quad \left(\frac{\alpha^i}{\alpha^i - 1}\right)^{\frac{1}{2}} \left(1 - \frac{1}{\alpha^\delta}\right) \frac{e^{-\frac{1}{12}\left(1+\frac{1}{\delta(\alpha^i-1)}\right)}}{\sqrt{2\pi}} \qquad \textit{and}$$

$$c_{\alpha,\delta,i} \quad \leq \quad \left(\frac{\alpha^i}{\alpha^i - 1}\right)^{\frac{3}{2}} \left(1 - \frac{1}{\alpha^\delta}\right) \frac{e^{\frac{1}{12}}}{\sqrt{2\pi}}.$$

The maximum is attained by an s in $[\delta(\alpha^i - 1) - 1, \delta(\alpha^i - 1)]$ and this is the only maximum.

Proof. The positive real s_0 maximizing the quotient

$$\frac{\bar{p}_i(\delta, s)}{\bar{p}_R(\delta, s)} \quad = \quad \binom{\delta + s}{\delta} \left(1 - \frac{1}{\alpha^i}\right)^s \left(1 - \frac{1}{\alpha^\delta}\right) \alpha^{\delta(1-i)} \tag{3.11}$$

also maximizes the logarithm of this quotient. Writing the binomial coefficient as $(s + 1) \cdot \ldots \cdot (s + \delta)/\delta!$ we obtain

$$\frac{\partial}{\partial s}\left[\log \frac{\bar{p}_i(\delta, s)}{\bar{p}_r(\delta, s)}\right] = \sum_{j=1}^{\delta} \frac{1}{s + j} \; + \; \log\left(1 - \frac{1}{\alpha^i}\right).$$

The second derivative is negative, thus we obtain the maximum as the solution of

$$0 = \sum_{j=1}^{\delta} \frac{1}{s + j} \; + \; \log\left(1 - \frac{1}{\alpha^i}\right),$$

which is equivalent to

$$\log \frac{\alpha^i}{\alpha^i - 1} \quad = \quad \sum_{j=1}^{\delta} \frac{1}{s + j}. \tag{3.12}$$

By approximating the sum by integrals, we get

$$\log(s + \delta + 1) - \log(s + 1) \;\leq\; \sum_{j=1}^{\delta} \frac{1}{s + j} \;\leq\; \log(s + \delta) - \log s. \tag{3.13}$$

By inserting equation (3.12) into inequalities (3.13) and applying the exponential function, we obtain that the solution s_0 fulfills

$$\frac{s+\delta+1}{s+1} \;\le\; \frac{\alpha^i}{\alpha^i-1} \;\le\; \frac{s+\delta}{s} \tag{3.14}$$

which is equivalent to

$$\delta(\alpha^i-1)-1 \;\le\; s \;\le\; \delta(\alpha^i-1).$$

Recall that by Stirling's approximation [124]

$$n! \;=\; \sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n}\, r_n$$

with $e^{\frac{1}{12n+1}} < r_n < e^{\frac{1}{12n}}$. Inserting the upper bound of $\delta(\alpha^i-1)$ on s_0 in equation (3.11), Stirling's approximation yields

$$\binom{\delta+s_0}{\delta} \;\le\; \binom{\alpha^i\delta}{\delta} \;=\; \frac{1}{\sqrt{2\pi}}\,\delta^{-\frac{1}{2}}\left(\frac{\alpha^i}{\alpha^i-1}\right)^{(\alpha^i-1)\delta+\frac{1}{2}}\alpha^{\delta i}\, r_{\alpha,\delta,i} \tag{3.15}$$

with $r_{\alpha,\delta,i} = \frac{r_{\alpha^i\delta}}{r_\delta\, r_{(\alpha^i-1)\delta}}$.

Thus,

$$r_{\alpha,\delta,i} \;<\; \frac{e^{\frac{1}{12\alpha^i\delta}}}{e^{\frac{1}{12\delta+1}}\cdot e^{\frac{1}{12(\alpha^i-1)\delta+1}}} \;\le\; \frac{e^{\frac{1}{12}}}{1\cdot 1}$$

and

$$r_{\alpha,\delta,i} \;>\; \frac{e^{\frac{1}{12\alpha^i\delta+1}}}{e^{\frac{1}{12\delta}}\cdot e^{\frac{1}{12(\alpha^i-1)\delta}}} \;\ge\; \frac{1}{e^{\frac{1}{12}}\cdot e^{\frac{1}{12(\alpha^i-1)\delta}}} \;=\; e^{-\frac{1}{12}\left(1+\frac{1}{\delta(\alpha^i-1)}\right)}.$$

For the upper bound in the lemma, we insert the bound 3.15 on the binomial coefficient at s_0 into equation 3.11. The remaining part of the right hand side of equation 3.11 is monotone decreasing in s and we therefore insert the lower bound $\delta(\alpha^i-1)-1$ on s_0 from equation 3.14 for this part. This yields

$$\begin{aligned}
&\frac{\bar{p}_i(\delta,s)}{\bar{p}_R(\delta,s)}\\
&<\; \frac{1}{\sqrt{2\pi}}\delta^{-\frac{1}{2}}\left(\frac{\alpha^i}{\alpha^i-1}\right)^{\delta(\alpha^i-1)+\frac{1}{2}}\alpha^{\delta i} r_{\alpha,\delta,i}\left(\frac{\alpha^i-1}{\alpha^i}\right)^{\delta(\alpha^i-1)-1}\left(1-\frac{1}{\alpha^\delta}\right)\alpha^{\delta(1-i)}\\
&\le\; \frac{\alpha^\delta}{\delta^{\frac{1}{2}}}\left(\frac{\alpha^i}{\alpha^i-1}\right)^{\frac{3}{2}}\left(1-\frac{1}{\alpha^\delta}\right)\frac{e^{\frac{1}{12}}}{\sqrt{2\pi}}.
\end{aligned}$$

For the lower bound in the lemma we insert $s_0' = \delta(\alpha^i-1)$ into equation 3.11. This gives

$$\begin{aligned}
&\frac{\bar{p}_i(\delta,s_0')}{\bar{p}_R(\delta,s_0')}\\
&=\; \frac{1}{\sqrt{2\pi}}\delta^{-\frac{1}{2}}\left(\frac{\alpha^i}{\alpha^i-1}\right)^{\delta(\alpha^i-1)+\frac{1}{2}}\alpha^{\delta i} r_{\alpha,\delta,i}\left(\frac{\alpha^i-1}{\alpha^i}\right)^{\delta(\alpha^i-1)}\left(1-\frac{1}{\alpha^\delta}\right)\alpha^{\delta(1-i)}\\
&\le\; \frac{\alpha^\delta}{\delta^{\frac{1}{2}}}\left(\frac{\alpha^i}{\alpha^i-1}\right)^{\frac{1}{2}}\left(1-\frac{1}{\alpha^\delta}\right)\frac{e^{-\frac{1}{12}\left(1+\frac{1}{\delta(\alpha^i-1)}\right)}}{\sqrt{2\pi}}.
\end{aligned}$$

□

Conflicts of a Round. Next we bound for a configuration the probability of appearing in a round weighted by its number of stoppers. We do this for an arbitrary round but the interesting case is the last round, i.e., the case $i = 1$.

Lemma 3.12. *For $\alpha > 1$, $\delta \geq 1$, and $i \geq 1$ the following holds*

$$\begin{aligned}\delta(\alpha^i - 1) \max_{s>0} \frac{\bar{p}_i(\delta, s)}{\bar{p}_R(\delta, s)} &\leq \max_{s>0} \frac{s\,\bar{p}_i(\delta, s)}{\bar{p}_R(\delta, s)} \\ &\leq ((\delta + 1)(\alpha^i - 1) + 1) \max_{s>0} \frac{\bar{p}_i(\delta, s)}{\bar{p}_R(\delta, s)}.\end{aligned}$$

Proof. If

- s_0 maximizes $\frac{\bar{p}_i(\delta,s)}{\bar{p}_R(\delta,s)}$ and
- s_1 maximizes $s\,\frac{\bar{p}_i(\delta,s)}{\bar{p}_R(\delta,s)}$

then

$$s_0 \frac{\bar{p}_i(\delta, s_0)}{\bar{p}_R(\delta, s_0)} \leq s_1 \frac{\bar{p}_i(\delta, s_1)}{\bar{p}_R(\delta, s_1)} \leq s_1 \frac{\bar{p}_i(\delta, s_0)}{\bar{p}_R(\delta, s_0)}.$$

Thus, it suffices to prove that $s_1 \leq (\delta + 1)(\alpha^i - 1) + 1$. We get s_1 as solution of

$$0 = \frac{\partial}{\partial s} \log \frac{s\,\bar{p}_i(\delta, s)}{\bar{p}_R(\delta, s)} = \frac{1}{s} + \frac{\partial}{\partial s} \log \frac{\bar{p}_i(\delta, s)}{\bar{p}_R(\delta, s)} = \sum_{j=0}^{\delta} \frac{1}{s+j} + \log \frac{\alpha^i - 1}{\alpha^i},$$

which is equivalent to

$$\log \frac{\alpha^i}{\alpha^i - 1} = \sum_{j=0}^{\delta} \frac{1}{s+j}.$$

As in the proof of Lemma 3.11 we approximate the sum by integrals which yields

$$\frac{s + \delta + 1}{s} \leq \frac{\alpha^i}{\alpha^i - 1} \leq \frac{s + \delta}{s - 1}.$$

This is equivalent to

$$(\delta + 1)(\alpha^i - 1) \leq s \leq (\delta + 1)(\alpha^i - 1) + 1.$$

□

Using Lemma 3.12 we can bound $\mathrm{E}_{BRIO^*}\left[C_\ell^1(n)\right]$ as follows. Recall that $C_\ell^1(n)$ is the sum of conflict sizes of configurations with all stoppers in the last round. If we only sum over the configurations of the last round then Lemma 3.12 directly yields a bound on the expected sum of conflict sizes. Now consider the probability $p_1^*(\delta, s)$ that a configuration with δ triggers and s stoppers has all its stoppers in the last round. Since the probability for a point to be in the last round is $(1 - 1/\alpha)$, we have

$$p_1^*(\delta, s) = \left(1 - \frac{1}{\alpha}\right)^s = p_1(\delta, s)\left(1 - \frac{1}{\alpha^\delta}\right)^{-1}.$$

From this we directly get the following observation which completes the proof of Theorem 3.9. By $\bar{p}_1^*(\delta, s)$ we denote the real function corresponding to the probability $p_1^*(\delta, s)$.

Observation 3.13. *For $\alpha > 1$, $\delta \geq 1$, and $i \geq 1$ the following holds*

$$\max_{s>0} \frac{s\,\bar{p}_1^*(\delta, s)}{\bar{p}_R(\delta, s)} = \max_{s>0} \frac{s\,\bar{p}_1(\delta, s)}{\bar{p}_R(\delta, s)} \left(1 - \frac{1}{\alpha^\delta}\right)^{-1}.$$

Probability of Appearing in the Construction. We want to bound for a configuration the probability of appearing in a construction con BRIO relative to the probability of appearing in a randomized incremental construction. Summing up the probabilities obtained in Lemma 3.11 for the individual rounds does not yield a useful bound. The problem is that the values of s maximizing the quotient of the two probabilities is different for the individual rounds. Instead we compute a bound for the same value of s for all rounds.

Lemma 3.14. *For $\alpha \geq 1 + \frac{1}{\delta}$ and $\delta \geq 1$ the following holds*

$$\max_{n \geq s > 0} \frac{\bar{p}_B(\delta, s)}{\bar{p}_R(\delta, s)} \leq \frac{\alpha^\delta}{\sqrt{\delta}} \cdot \frac{e^{13/12}}{\sqrt{2\pi}} \cdot \left(1 - \frac{1}{\alpha^\delta}\right) \cdot \frac{\alpha+1}{\alpha-1} \cdot \frac{\delta\alpha - 1}{\delta(\alpha-1) - 1}.$$

Proof. Let β be defined by $s = (\beta - 1)\delta$. Then,

$$\frac{\bar{p}_B(\delta, s)}{\bar{p}_R(\delta, s)} = \binom{\beta\delta}{\delta} \sum_{i=1}^{\infty} \left(\frac{\alpha^i - 1}{\alpha^i}\right)^{(\beta-1)\delta} \left(1 - \frac{1}{\alpha^\delta}\right) \alpha^{\delta(1-i)}.$$

By Stirling's approximation (cf. inequality 3.15),

$$\binom{\beta\delta}{\delta} \leq \frac{1}{\sqrt{2\pi}} \delta^{-\frac{1}{2}} \left(\frac{\beta}{\beta-1}\right)^{(\beta-1)\delta + \frac{1}{2}} \beta^\delta e^{\frac{1}{12}}.$$

Therefore,

$$\frac{\bar{p}_B(\delta, s)}{\bar{p}_R(\delta, s)} \leq \frac{\alpha^\delta}{\sqrt{\delta}} \left(1 - \frac{1}{\alpha^\delta}\right) \frac{e^{1/12}}{\sqrt{2\pi}} \frac{\beta}{\beta-1} \sum_{i=1}^{\infty} \left(\frac{\beta}{\alpha^i} \left(\frac{\beta}{\beta-1} \cdot \frac{\alpha^i - 1}{\alpha^i}\right)^{\beta-1}\right)^\delta$$

We now bound the sum

$$\sum_{i=1}^{\infty} \left(\frac{\beta}{\alpha^i} \left(\frac{\beta}{\beta-1} \cdot \frac{\alpha^i - 1}{\alpha^i}\right)^{\beta-1}\right)^\delta.$$

Each summand can be bounded by

$$\begin{aligned} \frac{\beta}{\alpha^i} \left(\frac{\beta}{\beta-1} \cdot \frac{\alpha^i - 1}{\alpha^i}\right)^{\beta-1} &= \frac{\beta}{\alpha^i} \left(1 - \left(\frac{\beta - \alpha^i}{(\beta-1)\alpha^i}\right)\right)^{\beta-1} \\ \leq \frac{\beta}{\alpha^i} e^{-\frac{\beta-\alpha^i}{\alpha^i}} &= e\frac{\beta}{\alpha^i} e^{-\frac{\beta}{\alpha^i}} \leq e\frac{\alpha^i}{\beta}. \end{aligned} \tag{3.16}$$

For the first inequality we use that $1 + x \leq e^x$ for all $x \in \mathbb{R}$. For the second inequality we use $xe^{-x} = x^2 e^{-x} \frac{1}{x} \leq \frac{1}{x}$.

Each summand is maximized for $\beta = \alpha^i$. To see this, we take the derivative in β of the summand

$$f_{\alpha^i}(\beta) := \frac{\beta}{\alpha^i}\left(\frac{\beta}{\beta-1}\cdot\frac{\alpha^i-1}{\alpha^i}\right)^{\beta-1}$$

and set it to zero.

$$\begin{aligned}\frac{d}{d\beta}f_{\alpha^i}(\beta) &= \left(\log\frac{\beta}{\alpha^i}+1\right)f_{\alpha^i}(\beta)+\left(\log\frac{\alpha^i-1}{\beta-1}-1\right)f_{\alpha^i}(\beta)\\ &= \log\frac{\beta}{\beta-1}\frac{\alpha^i-1}{\alpha^i}f_{\alpha^i}(\beta).\end{aligned}$$

Since $f_{\alpha^i}(\beta) > 0$ we get that $\frac{d}{d\beta}f_{\alpha^i}(\beta) = 0$ if $\log\frac{\beta}{\beta-1}\frac{\alpha^i-1}{\alpha^i} = 0$. This is the case if $\frac{\beta}{\beta-1}\frac{\alpha^i-1}{\alpha^i} = 1$ which is equivalent to $\beta = \alpha^i$. This is a maximum because $f_{\alpha^i}(\alpha^i) = 1$ and $f_{\alpha^i}(\beta)$ tends to $\frac{1}{\alpha^i} < 1$ for $\beta \to 1$ and $f_{\alpha^i}(\beta)$ tends to 0 for $\beta \to \infty$ by the bound (3.16).

Thus, since each summand is maximized for $\beta = \alpha^i$ and for this it is 1, each summand is at most 1. Ignoring the exponent δ only increases the sum.

Let $i_0 := \lceil \log_\alpha \beta \rceil$. We split the sum at i_0, i.e., into summands with $\alpha^i \geq \beta$ and $\alpha^i < \beta$, and handle these cases separately. For $\alpha^i \geq \beta$ we use

$$\left(\frac{\beta}{\beta-1}\right)^{\beta-1} \leq e \quad\text{and}\quad \frac{\alpha^i-1}{\alpha^i} < 1$$

and bound the partial sum by

$$\begin{aligned}&\sum_{i=i_0}^{\infty}\left(\frac{\beta}{\alpha^i}\left(\frac{\beta}{\beta-1}\cdot\frac{\alpha^i-1}{\alpha^i}\right)^{\beta-1}\right)^{\delta}\\ &\leq e\sum_{i=i_0}^{\infty}\frac{\beta}{\alpha^i} = e\sum_{i=0}^{\infty}\frac{\beta}{\alpha^{i_0}}\alpha^{-i} \leq e\sum_{i=0}^{\infty}\alpha^{-i} = e\frac{\alpha}{\alpha-1}.\end{aligned}$$

For $\alpha^i \leq \beta$ we use the estimate (3.16) to bound the partial sum by

$$\begin{aligned}&\sum_{i=1}^{i_0-1}\left(\frac{\beta}{\alpha^i}\left(\frac{\beta}{\beta-1}\cdot\frac{\alpha^i-1}{\alpha^i}\right)^{\beta-1}\right)^{\delta}\\ &\leq e\sum_{i=1}^{i_0-1}\frac{\beta}{\alpha^i} \leq \frac{e}{\alpha^{i_0}}\sum_{i=1}^{i_0-1}\alpha^i = \frac{e}{\alpha^{i_0}}\cdot\frac{\alpha_{i_0}-\alpha}{\alpha-1} \leq \frac{e}{\alpha-1}.\end{aligned}$$

Thus, we can bound the whole sum by $e\frac{\alpha+1}{\alpha-1}$.

It remains to bound $\frac{\beta}{\beta-1} = 1 + \frac{1}{\beta-1}$. We bound this term for β_2 for which $s_2 = \delta(\beta_2 - 1)$ maximizes $\frac{\bar{p}_B(\delta,s)}{\bar{p}_R(\delta,s)}$ for given α and δ. By Lemma 3.11 for every $i \geq 1$ the quotient $\frac{\bar{p}_i(\delta,s)}{\bar{p}_R(\delta,s)}$ is maximized for an s with

$$s \;\geq\; \delta(\alpha^{-i}-1)-1 \;\geq\; \delta(\alpha-1)-1$$

and $\frac{\bar{p}_i(\delta,s)}{\bar{p}_R(\delta,s)}$ is monotone increasing for smaller s. Therefore, also $s_2 > \delta(\alpha-1)-1$ and since $\alpha > 1 + \frac{1}{\delta}$ we have $\delta(\alpha-1)-1 > 0$. For β_2 this yields

$$\delta(\beta_2 - 1) \ \geq \ \delta(\alpha - 1) - 1 \ > \ 0,$$

and therefore

$$\frac{\beta_2}{\beta_2 - 1} \ \leq \ \frac{\delta\alpha - 1}{\delta(\alpha - 1) - 1}.$$

□

3.4 Expected-Case Analysis for Random Points

In Section 3.3 we bounded the expected running time of incremental constructions con BRIO for an arbitrary order within the rounds. In this section we bound the expected running time for a special class of orders within rounds, i.e., space-filling curve orders (see Chapter 2). We formulate the algorithm for a set of points in $\mathbb{R}^d$ (Algorithm 5) but in this section we will analyze the algorithm for points in two dimensions drawn independently and uniformly at random from a bounded convex region. The analysis also holds for pseudo-uniformly distributed points, i.e., points with a density function bounded from above and below by positive constants. This does not include normally distributed points which we will consider in the next chapter.

Algorithm 5: Incremental Construction along Space-Filling Curves

Input: Point set in $\mathbb{R}^d$
Output: Delaunay tessellation of the point set

1 Compute BRIO with SFC in rounds:
1.1 Sample points to rounds (using coin flips with sampling parameter $\alpha > 1$),
1.2 Order points in a round using a space-filling curve order, for every other round use reversed order.
2 Incrementally construct Delaunay tessellation using order from step 1: In each step do
2.1 Locate new point from the previously inserted point by walking,
2.2 Update Delaunay tessellation.

The details of the steps of the algorithm have been covered for step 1.1 in Sections 3.2.4 and 3.3, for step 1.2 in Section 2.2.2, and for step 2.1 and step 2.2 in Section 3.2.2.

In step 1.2 we assume that the bounding cube is chosen such that its side length either is the largest L_∞-distance of two points, or if a bounded region in which the points lie is known, it is the L_∞-diameter of the region. We reverse the space-filling curve order for every other round such that also for non-closed space-filling curves like the Hilbert curve the concatenated curve is continuous. The motivation for this is to have a short distance and therefore fewer walking steps between the last point of a round and the first point of the next round.

This is not important for the analysis since even if the cost of the first point location step in a round would be linear, this would not change the asymptotic bound on the expected running time.

3.4.1 Counting Intersections

To analyze the running time of the point location by walking it suffices to analyze the time required in the last round. This bound can then be applied to the other rounds with the number of points being a random variable depending on the round. Since we obtain a linear bound for one round we can further use linearity of expectation to compute the expected total linear running time.

Assume N points are distributed independently and uniformly at random in a bounded convex region C of area 1, where

- $N = m + n$ with
- n points already inserted in the Delaunay triangulation and
- m points to be inserted in the next round.

To insert the m points, a space-filling curve tour through these points is constructed.

The points are located by traversing the Delaunay triangulation along the tour. Therefore, the time for locating the points is proportional to the number of intersections between the tour and the Delaunay triangulation. We consider two variants. In the first variant, the points of the round are located along the space-filling curve without directly inserting them into the Delaunay triangulation. We bound the expected number of intersections for this case first and then extend it to the second variant. In the second variant, points are directly inserted. Thus, the triangulation changes while the points are located along the space-filling curve.

Exclusion Regions. The method used to obtain a bound on the number of intersections can be applied to other classes of triangulations for which an *exclusion region* can be defined: This is a region that can be placed around each edge in such a way that the region is empty on one of the sides of the edge. This is necessary (but not sufficient) for the edge to be in the triangulation. For the Delaunay triangulation the disc with the edge as diameter is an exclusion region.

With some weak assumptions on the exclusion region, Dickerson et al. [53] observed that possible edges of such a triangulation of uniformly distributed points are expected to be either short or near to the boundary. In our case this means that endpoints of Delaunay edges intersecting the tour are either near to the tour or near to the boundary of the bounded convex region C.

By Observation 2.1 a space-filling curve tour through m points in the unit square has length at most $c_T\sqrt{m}$ where c_T is the Hölder constant of the space-filling curve. The following lemma and proposition are formulated more generally for traveling salesperson problem heuristics with an expected tour length of $O(\sqrt{m})$ through m points in the unit square. This holds for any expected constant-factor approximation algorithm for the traveling salesperson problem.

Lemma 3.15. *Let $n + m$ points be distributed independently and uniformly in a bounded convex region C in the plane of area 1. Let the n points define a Delaunay triangulation. Let a tour through the remaining m points be given which depends only on the m points and has expected length bounded by $c_T\sqrt{m}$ for a constant c_T. The expected number of intersections between line segments of the tour that have distance at least $c_0\sqrt{\log n/n}$ from the boundary ∂C and edges of the triangulation is bounded by $c_1 c_T\sqrt{mn}+c_2 m$ where c_0 is an absolute constant and c_1, c_2 are constants depending on geometric properties of C.*

Proof. We consider only line segments of the tour which have distance at least $c_0\sqrt{\log n/n}$ from the boundary ∂C where the constant c_0 is the same as in Lemma 3.2. We call these segments far from the boundary. For a given line segment L of the tour far from the boundary Lemma 3.2 bounds the expected number of intersections between L and the Delaunay triangulation by $c_1\sqrt{n}+c_2$ for constants c_1 and c_2 depending on geometric properties of C. Therefore, for a tour through the m points of length $c_T\sqrt{m}$ the expected number of intersections between line segments of the tour far from the boundary and the Delaunay triangulation is bounded by $c_1\, c_T\, \sqrt{nm} + c_2\, m$.

This bound also applies to the case of a tour through random points if they are independent of the points of the triangulation. To see this, consider the random points $Y_1, \ldots, Y_m$ through which the tour is taken (not yet ordered). Let I be the number of intersections and let $|T|$ be the length of the tour. I is a random variable depending on the points of the tour and of the Delaunay triangulation. $|T|$ is a random variable depending only on the tour.

By the independence of the points of the tour and the Delaunay triangulation, we have

$$\mathrm{E}\,[I \mid Y_1 = y_1, \ldots, Y_m = y_m] \;\leq\; c_0\,\sqrt{n}\,\mathrm{E}\,[|T| \mid Y_1 = y_1, \ldots, Y_m = y_m] \;+\; c_1 m$$

and therefore

$$\begin{aligned}
\mathrm{E}\,[I] &= \mathrm{E}\,[\mathrm{E}\,[I \mid Y_1, \ldots, Y_m]] \\
&\leq \mathrm{E}\left[c_0\sqrt{n}\,\mathrm{E}\,[|T| \mid Y_1, \ldots, Y_m] + c_1 m\right] \\
&= c_0\sqrt{n}\,\mathrm{E}\,[|T|] + c_1 m \\
&\leq c_0 c_2\sqrt{nm} + c_1 m.
\end{aligned}$$

□

With Lemma 3.15 we now have all necessary ingredients for a linear expected time algorithm for constructing the Delaunay triangulation. For this, consider the following algorithm. Use an incremental construction con BRIO and in each round of n points inserted so far and m points to be inserted:

- Build in $O(n)$ time a point location data structure in the triangulation with $O(\log n)$ query time.

- Locate the points to be inserted along a traveling salesperson problem heuristic which can be computed in $O(m)$ time and has expected $O(\sqrt{m})$ length. As fallback use the point location data structure. For instance this can be done by running the point location schemes in parallel, i.e., one step of point location by walking , followed by one step of point location in the data structure, and so on.

- Insert the points using a history with the triangulation of the n points as first level.

A possible point location data structure is Kirkpatrick's point location hierarchy [85]. Points can be inserted during the walk if the walk is performed on the original triangulation, i.e., on the first level of the hierarchy.

Proposition 3.16. *The algorithm given above constructs the Delaunay triangulation of independently and uniformly distributed points in a bounded convex region in the plane in linear expected time.*

Proof. By Lemma 3.15, all points except those near the boundary can be located by walking in $O(\sqrt{mn} + m)$ time. The area of the boundary region is $O(\sqrt{\log n/n})$. Thus, the expected number of points near the boundary and therefore also the expected number of tour segments with an endpoint near the boundary are in $O(m\sqrt{\log n/n})$. For each of these tour segments the point location takes $O(\log n)$ time. Thus, in total the point location of all points of a round in the previous Delaunay triangulation takes $O(\sqrt{mn} + m + m\frac{\log^{3/2} n}{\sqrt{n}}) \subset O(m + n)$ time. The additional cost of using the history is linear by Corollary 3.10. □

However, we are interested in a linear expected time algorithm without point location data structure. We will prove that the fallback to a point location data structure for points near the boundary is not necessary. For this, we do an explicit boundary analysis. We will also show that the history is not needed.

3.4.2 Boundary Analysis

To treat segments near the boundary we first quantify what it means that the edges of the triangulation are likely to be short or near to the boundary in the following lemma. The proof of Lemma 3.17 generalizes and simplifies the proof for the case of the unit square [49].

For the purpose of the analysis in this chapter it would suffice to prove that a Delaunay edge cannot be too long if one of the endpoints is not too close to the boundary. Nonetheless, we prove the slightly stronger statement that a Delaunay edge cannot be too long if there is any point on the edge, i.e., not necessarily an endpoint, which is not too close to the boundary.

Lemma 3.17. *Let $n > 2$ points be distributed independently and uniformly in a bounded convex region C of area 1 in the plane. Denote by $D_{w,\ell}$ the event that the Delaunay triangulation of the points contains an edge which has a point on*

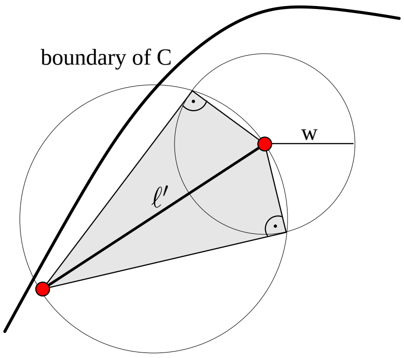


Figure 3.3: An exclusion region for a Delaunay edge with one endpoint far from the boundary of C.

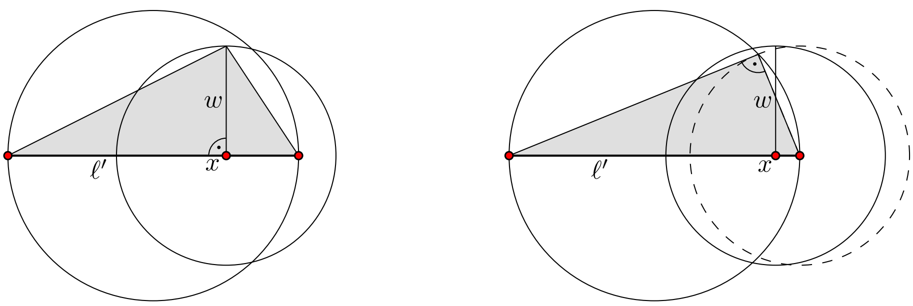
it with distance at least w to the boundary of C and which is longer than ℓ. For any $t > 1$ and $\ell \geq tw$

$$P[D_{w,\ell}] \leq n^2 e^{-(n-2)w\ell\sqrt{1-1/t^2}/2}.$$

In particular, if $\ell \geq 3w$ and $w\ell \geq 6\sqrt{2}\log n/(n-2)$ then

$$P[D_{w,\ell}] \leq 1/n^2.$$

Proof. First consider the case of an edge with length $\ell' \geq \ell$ and an endpoint with distance more than w to the boundary of C as shown in Figure 3.3. The circle with the edge as diameter and the circle of radius w around the endpoint with distance at least w to the boundary intersect since $w < \ell$. The intersection points of the circles lie in C since the circle of radius w lies in C. The edge together with these two intersection points form two right triangles contained in C and these give an exclusion region for the edge. The area of one of these triangles is bounded from below by $1/2 \cdot w\sqrt{\ell^2 - w^2} \geq w\ell\sqrt{1-1/t^2}/2 =: a$. Therefore, the probability that the triangle is empty is bounded from above by $(1-a)^{n-2}$. The probability that both triangles of an edge are empty is bounded by $2(1-a)^{n-2}$.



(a) (b)

Figure 3.4: Cases if a point on the edge has distance w to the boundary.

Next consider the situation that both endpoints of the edge are closer than w to the boundary but a point x on the edge has distance at least w to the boundary. There are two cases as shown in Figure 3.4. The first case is shown in Figure 3.4(a): The circle with the Delaunay edge as diameter contains the line segments of length w starting at x and perpendicular to the edge. Then the resulting triangles both have area $w\ell/2 > w\ell\sqrt{1-1/t^2}/2$ and form an exclusion region. The second case is shown in Figure 3.4(b): The circle with the Delaunay edge as diameter does not contain these segments. Then the intersections of this circle and the circle with radius w around the endpoint closer to x (dashed circle) are in the circle around x. They are therefore in C and we can take the same exclusion region as in the case where this endpoint has distance w to the boundary.

There are $\binom{n}{2}$ possible edges and thus

$$\mathrm{P}\left[D_{w,\ell}\right] \;\leq\; \binom{n}{2} 2(1-a)^{n-2} \;\leq\; n^2 e^{-(n-2)a} \;=\; n^2 e^{-(n-2)w\ell\sqrt{1-1/t^2}/2}.$$

In particular, for $t = 3$ and $w\ell \geq 6\sqrt{2}\log n/(n-2)$ we get

$$\begin{aligned}\mathrm{P}\left[D_{w,\ell}\right] &\leq n^2 e^{-(n-2)(6\sqrt{2}\log n/(n-2))\sqrt{1-1/3^2}/2} \\ &= n^2 e^{-4\log n} = \frac{1}{n^2}.\end{aligned}$$

□

Let us note that the exclusion region was chosen with w much smaller than ℓ in mind. Better constants, in particular for w close to ℓ, can be obtained by including the whole intersection of the disk with radius w around the one endpoint and the disk with the edge as diameter into the exclusion region. We illustrate this by the following remark.

Remark 3.18. *With the same setting as in Lemma 3.17 but with $w = \ell$, the probability $P[D_{w,\ell}]$ can be bounded from above by $n^2 e^{-(n-2)w\ell\pi/8}$.*

Proof. For $w = \ell$ the circle of radius w contains a circle of diameter ℓ halved by the Delaunay edge of length $\ell' \geq \ell$. This gives an exclusion region for which each side as area

$$a := \frac{\ell^2}{8}\pi = \frac{w\ell}{8}\pi.$$

Thus,

$$\mathrm{P}\left[D_{w,\ell}\right] \;\leq\; \binom{n}{2} 2(1-a)^{n-2} \;\leq\; n^2 e^{-(n-2)a} \;=\; n^2 e^{-(n-2)w\ell\pi/8}.$$

□

This gives us a bound on the number of Delaunay edges that can intersect the line segments of the tour:

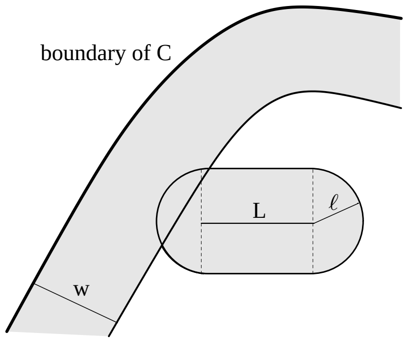


Figure 3.5: Area for endpoints of Delaunay edges intersecting L.

Lemma 3.19. *The expected number of intersections of a Delaunay triangulation of n points and a tour along a Hölder-$1/2$ space-filling curve through m points where the $m+n$ points are distributed independently and uniformly in a bounded convex region is in $O(m+\sqrt{mn})$ and therefore linear in the total number of points.*

Proof. For line segments of the tour not near the boundary, Lemma 3.15 gives a bound of $O(m+\sqrt{mn})$ on the number of intersections with the Delaunay triangulation. Next we consider the line segments of the tour near the boundary. Without loss of generality, the area of C is 1.

The expected number of points near the boundary is bounded by

$$m' := c_0\,|\partial C|\,m\,\sqrt{\log n/n}.$$

Each of these points can be adjacent to two line segments resulting in an upper bound of $k := 2m'$ on the expected number of line segments with at least one point near the boundary. Each of these line segments can contribute one further endpoint resulting in a bound of $3m'$ on the expected number of endpoints of such line segments. These line segments are also part of a space-filling curve tour through only these endpoints. Therefore, by Observation 2.1 and Jensen's inequality the expected total length of these line segments is bounded by $\ell_\Sigma := c_T\sqrt{3m'}$ for a constant c_T depending on the space-filling curve.

For $n > 2$ we choose

$$\begin{aligned} w &:= (\log n/(n-2))^{2/3} \quad \text{and} \\ \ell &:= 6\sqrt{2}(\log n/(n-2))^{1/3}. \end{aligned}$$

Then for $n > 2$, it is $\ell \geq 3w$ and $w\ell \geq 6\sqrt{2}\log n/(n-2)$. Thus, we have by Lemma 3.17, that with probability at least $1-\frac{1}{n^2}$ line segments are closer than w to the boundary or shorter than l. If this event occurs, only Delaunay edges with endpoints which have distance at most ℓ to one of the line segments of the tour, or distance at most w from the boundary can intersect the space-filling curve tour. If this event does not occur, then in the worst case a tour segment is intersected by all edges of the triangulation, which gives $O(1/n)$ intersections in expectation. In total, this contributes at most $O(m/n)$ expected intersections.

In the case that the event occurs, Figure 3.5 shows for a single line segment L the region in which both endpoints of an intersecting Delaunay edge must lie. The area of this region is bounded from above by $|\partial C|w+\pi\ell^2+2\ell|L|$. Therefore the expected number of endpoints of edges that intersect a line segment L is bounded by $n(|\partial C|w+\pi\ell^2+2\ell|L|)$. Because of planarity there are at most three times that many edges intersecting L. For k line segments of total expected length L_Σ this yields a bound of

$$3n(|\partial C|kw + \pi\ell^2 k + 2\ell L_\Sigma)$$

on the number of intersecting edges. Inserting the values for w, ℓ, k, and ℓ_Σ gives that the number of intersections near the boundary can be bounded by

$$\begin{aligned} &3n(|\partial C|2m'w + \pi\ell^2 2m' + 2\ell c_T\sqrt{3m'}) \\ \le\ & 6nm'(\log n/n)^{2/3}(|\partial C| + \pi) + 18c_T\sqrt{6}nm'^{1/2}(\log n/n)^{1/3} \\ \in\ & O\left(m\frac{\log^{7/6} n}{n^{1/6}} + m^{1/2}n^{5/12}\log^{7/12} n\right) \subset O(m+\sqrt{mn}). \end{aligned}$$

Adding up the bounds for segments near the boundary and not near the boundary gives the bound of the lemma. □

As a direct consequence of Lemma 3.19 (together with Corollary 3.10), we see that we do not need an additional point location data structure for points near the boundary for a linear expected time algorithm for constructing the Delaunay triangulation.

Proposition 3.20. *Incremental constructions con BRIO using an order along a Hölder-1/2 continuous space-filling curve to locate points in the Delaunay triangulation of the previous round and using a history for further point location runs in linear expected time.*

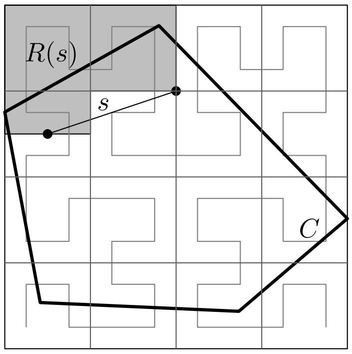
The algorithm in the proposition still uses the history within a round. In the following, we prove that this is not necessary.

3.4.3 Inserting Points while Walking

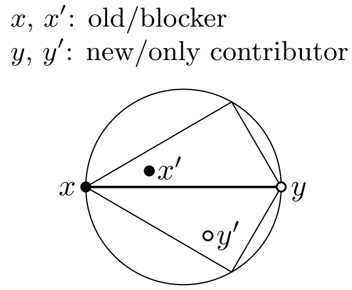
So far we have only considered point location in the Delaunay triangulation of the points of previous rounds. We now discuss the case where points are inserted during the traversal. That is, during the round walking is performed in the Delaunay triangulation into which some points of the current round have already been inserted.

As before, we restrict our attention to the last round. Let P_1 be the set of points before the last round, P_2 the set of points to be inserted in the last round. Let $n := |P_1|$ andd $m := |P_2|$. Let C be the region from which the points are drawn. If we consider a single line segment of the tour, we can bound the number of intersections with it by the following two observations.

First, the remaining points of the tour are not independent of this segment. The tour segment determines a part of C in which no other points lie. This is



(a) Tour segment and corresponding region



(b) Contributors and blockers

Figure 3.6: Illustration of density bound for tour segments and for the roles of points in the analysis of the Delaunay triangulation.

illustrated in Figure 3.6(a). If t_1 and t_2 are the preimages of the tour segment s selected by the space-filling curve heuristic with $t_1 \leq t_2$ and $R(s) := \psi([t_1, t_2])$ (shown in gray in the figure) then no other point of P_2 lies in the region $C \cap R(s)$. Conditional on the fact that $t_1 t_2$ is a tour segment, the remaining points are independently and uniformly distributed in $C_s := C \setminus R(s)$. Their density can be bounded from above if we assume the space-filling curve ψ to be *bi-measure preserving*, i.e., $\mu_1(A) = \mu_2(\psi(A))$ for any Borel set $A \in [0,1]$ and $\mu_1(\psi^{-1}(B)) = \mu_2(B)$ for any Borel set $B \in [0,1]^2$ where μ_1 and μ_2 denote the 1- and 2-dimensional measure, respectively. This allows us to work on uniform distributions on the preimage in the unit interval and the image in the unit square interchangeably [65]. Therefore we can bound the density of the point distribution by the density of the distribution of the preimages. We give the bound on the point density in Lemma 3.22.

Second, in the analysis of the boundary case (Lemmas 3.17 and 3.19) as well as the non-boundary case (Lemma 3.2) the points of the Delaunay triangulation occur in two different roles. On the one hand, points of $P_1 \cup P_2$ may contribute to the number of intersections as an endpoint of an intersecting edge. On the other hand, points of P_1 may also block potential edges between points of $P_1 \cup P_2$ because they lie in their exclusion region. We analyze the case with all points except the two endpoints of the current tour segment as contributors, i.e., possible endpoints of edges in the triangulation, but with only the points of P_1 as blockers.

For an illustration of the concept of contributors and blockers see Figure 3.6(b): Assume all four points are contributors but only x and x' are blockers, i.e., $x, x' \in P_1$ and $y, y' \in P_2$. If the edge between x and y is in the triangulation then at least one side of the exclusion region must be empty, i.e., one of the triangles, is "empty". The upper triangle is not empty because it contains x', and x' has been inserted in a previous round and is therefore a blocker. But the lower triangle is considered to be empty since y' is not considered to be a blocker. If y' is inserted before y then the edge xy would not actually appear. Even if one exclusion region is empty, the edge does not necessarily appear. We include all these (potential) edges in the analysis.

We will extend Lemmas 3.17, 3.2, and 3.19 to the case of contributors and blockers by Lemmas 3.23, 3.25, and 3.26, respectively. For the proof of Lemma 3.25 we will also extend Remark 3.18 to Remark 3.24. By this we handle the case of a Delaunay triangulation of a set of n points together with an arbitrary (i.e., in particular not necessarily independent) subset of a set of m points where the $n+m$ points are independent, and the n points are uniformly distributed in a bounded convex region C and the m points are uniformly distributed in an unknown but sufficiently large subset of C.

Let C be a convex region in $[0,1]^2$. We use μ_C to denote the area of C. Let $\psi\colon [0,1] \to [0,1]^2$ be a bi-measure preserving space-filling curve. For a point $y \in C$ let $\psi^*(y)$ be the preimage of y selected by the space-filling curve heuristic. Furthermore, for a line segment $s = (y, y')$ contained in C let

$$\mu_s := \mu\left([\psi^*(y), \psi^*(y')] \cap \psi^{-1}(C)\right)$$

be the length of the intersection of the interval corresponding to s and the preimage of C.

Observation 3.21. *Let m points be distributed independently and uniformly at random in a bounded convex region $C \subseteq [0,1]^2$. Let $\psi\colon [0,1] \to [0,1]^2$ be a bi-measure preserving space-filling curve. Conditional on the fact that s is a tour segment, the points on the tour except for the endpoints of the segment are distributed independently and uniformly in a region $C_s \subseteq C$ depending on s of measure $\mu_C - \mu_s$.*

In the following lemma we bound the probability that the space-filling curve tour contains a segment with large μ_s. This also shows that the measure of C_s is close to the measure of C and therefore the density of the uniform distribution on C_s close to the density of the distribution on C, i.e., close to $1/\mu_C$.

Lemma 3.22. *Let m points be distributed independently and uniformly at random in a convex region $C \subseteq [0,1]^2$. Let $\psi\colon [0,1] \to [0,1]^2$ be a bi-measure preserving space-filling curve. The probability of the event that a tour along ψ contains a segment s with $\mu_s > 2\frac{\log m}{m-1}\mu_C$ is bounded from above by $\frac{1}{m}$. Conditional on this event, the density of C_s is bounded from above by $\frac{1}{\mu_C}(1 + O(\log m/m))$.*

Proof. For a random starting point Y of a tour segment, let $A_t(Y)$ denote the event that an interval $[\psi^*(Y), b_t]$ starting at $\psi^*(Y)$ and with $\mu([\psi^*(Y), b_t] \cap C) \geq t\,\mu_C$ and $b_t \leq 1$ is empty. Then,

$$\mathrm{P}\left[A_t\right] \;\leq\; (1-t)^{m-1} \;\leq\; e^{-t(m-1)}.$$

Therefore,

$$\mathrm{P}\left[A_{2\log m/(m-1)}\right] \;\leq\; \left(1 - 2\frac{\log m}{m-2}\right)^{m-2} \;\leq\; e^{-2\frac{\log m}{m-2}(m-2)} \;=\; m^{-2}.$$

Summing up over all m possible starting points yields that the probability that the tour has a segment s with $\mu_s > 2\frac{\log m}{m-1}\mu_C$ is bounded by $\frac{1}{m}$.

If $\mu_s \le 2\frac{\log m}{m-1}$ holds for a line segment s then the measure of the remaining region is

$$\mu_C - \mu_s \;\ge\; \mu_C(1 - 2\frac{\log m}{m-1})$$

and the density function is bounded from above by

$$\frac{1}{\mu_C\left(1 - 2\frac{\log m}{m-1}\right)} = \frac{1}{\mu_C}\left(1 + 2\frac{\log m}{m-1}\left(1 - 2\frac{\log m}{m-1}\right)\right) = \frac{1}{\mu_C}O\left(1 + \frac{\log m}{m}\right).$$

□

Lemma 3.23. *Let $n+m$ points be distributed independently with the first $n > 2$ points distributed uniformly in a bounded convex region C of area 1 and the other m points arbitrarily distributed. Consider the Delaunay triangulation of the first n points and of an arbitrary subset of the other m points. Denote by $D_{w,\ell}$ the event that this contains an edge which has a point on it with distance at least w to the boundary of C and which is longer than ℓ. For any $t > 1$ and $\ell \ge tw$*

$$P[D_{w,\ell}] \;\le\; (n+m)^2 e^{-(n-2)w\ell\sqrt{1-1/t^2}/2}.$$

In particular, if $\ell \ge 3w$ and $w\ell \ge 6\sqrt{2}\log(n+m)/(n-2)$, then $P[D_{w,\ell}] \le 1/(n+m)^2$.

Proof. For a single point the bound remains the same as obtained in the proof of Lemma 3.17. Instead of $\binom{n}{2} \le n^2/2$ possible edges, there are now $\binom{n+m}{2} \le (n+m)^2/2$ possible edges which yields the bound on $\mathrm{P}[D_{w,l}]$. □

Remark 3.24. *With the same setting as in Lemma 3.17 but with $w = \ell$, the probability $P[D_{w,\ell}]$ can be bounded by $(n+m)^2 e^{-(n-2)w\ell\pi/8}$.*

Next we turn to the proof of an extended version of Lemma 3.2. A bound typically used in the probabilistic analysis of Delaunay triangulations to bound the length of Delaunay edges (and which in this chapter we have used so far only implicitly by using Lemma 3.2) is the following: Let P be a point set in the plane and $x \in P$. Let the plane be subdivided into sufficiently many cones originating at x. Assume there is a point of P with distance at most t to x in the interior of every cone. Then there is no edge in the Delaunay triangulation of P with x as an endpoint and longer than ct where $c > 0$ is a constant depending on the number of cones. A bound on c can be computed using exclusion regions. For instance, Katajainen and Koppinen [84] use in the analysis of their divide-and-conquer algorithm 16 cones which allows to choose $c = \sqrt{2}$.

Better bounds can be obtained by considering the *tentative Voronoi cell* of a point x. Bentley, Weide and Yao [12] use this argument with 8 cones (without explicitly computing c). We give the argument for 6 cones (see Figure 3.7): If we have a point in every cone then x together with the closest point to x in every cone (shown as white points in the figure) defines the tentative Voronoi cell of x, i.e., the Voronoi cell of x in the Voronoi diagram of these points. Now, a further point can only share a Voronoi edge (and therefore a Delaunay edge)

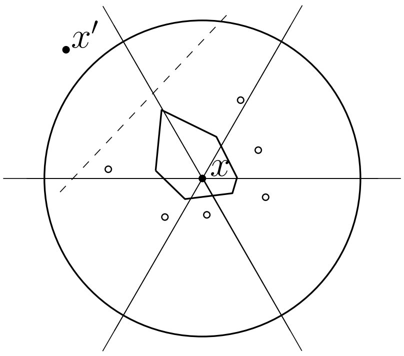


Figure 3.7: Tentative Voronoi cell of a point x.

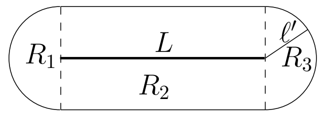


Figure 3.8: Points with distance ℓ' to a line segment.

with x if the bisector of this point and x intersects the tentative Voronoi cell of x. In the example of the figure, x' cannot share a Voronoi edge with x since the bisector of x and x' indicated by the dashed line does not intersect the tentative Voronoi cell. Since the angle formed by two points in the interior of neighboring cones and with x as a base is less than $3\pi/2$, the distance of x to any point in the tentative Voronoi cell is at most the distance to one of the closest neighbors in the cones. Thus, we have $c = 2$. In the figure, the corresponding circle is shown, i.e., no point outside the circle can have a Voronoi edge with x in common.

For the analysis of the Delaunay triangulation of uniformly distributed points in a bounded convex area such a bound can be used to bound the maximum edge length in a Delaunay triangulation at a point x as long as x is sufficiently far away from the boundary. As long as $t/2$ is less or equal to the distance of x to the boundary the probability $p(t)$ that the maximum edge length at x is larger than t is bounded by

$$p(t) \;\leq\; 6(1 - \pi(t/2)^2/6)^{n-1} \;=\; 6(1 - \pi t^2/24)^{n-1} \tag{3.17}$$

for n blocking points with density function 1. This bound does not depend on the number of contributing points. We use this bound together with Remark 3.24 to extend Lemma 3.2.

Lemma 3.25. *Let $n+m$ points be distributed independently with the first $n > 2$ points distributed uniformly in a bounded convex region C of area 1. Let L be a random line segment independent of the first n points. Let the conditional distribution of the remaining m points conditioned under the line segment have a density function bounded by $\lambda_m \geq 1$. Let DT be the Delaunay triangulation of the first n points and of an arbitrary subset of the other m points.*

If the distance of L to the boundary of C is more than $c_0\sqrt{\log(n+m)/(n-2)}$ then the expected number of intersections between DT and L is bounded by

$$c_1 \frac{n + \lambda_m m}{n} + c_2 |L| \frac{n + \lambda_m m}{\sqrt{n}},$$

where the constants can be chosen as $c_0 = 4\sqrt{6/\pi}$, $c_1 = 435$, and $c_2 = 72\sqrt{6}$.

Proof. We choose

$$w \;:=\; \sqrt{24/\pi}\sqrt{\log(n+m)/(n-2)} \quad \text{and} \quad \ell \;:=\; w.$$

By Remark 3.24 we have with probability at least $1-\frac{1}{n+m}$ that the event $D^c_{w,\ell}$ occurs, i.e., Delaunay edges with points with distance more than w to the boundary can have length at most ℓ. In the case of this event we have that if L has distance more than $c_0\sqrt{\log{(n+m)}/(n-2)} = 2\ell > w$ to the boundary then it can only be reached by Delaunay edges of length at most ℓ. The endpoints of these edges then have distance more than $\ell = w$ to the boundary, and therefore can only be reached by edges of length ℓ.

Figure 3.8 shows the set of points with a distance $\ell' \le \ell$ to L. Using the bound 3.17, we can bound the probability that a point with distance ℓ' to L has a Delaunay edge intersecting L by $6(1-\pi\ell'^2/24)^{n-1}$. The density of the distribution of the first n points is 1 and therefore the probability for a single point to reach L is bounded by

$$2\pi\int_0^\ell t6(1-\pi t^2/24)^{n-1}dt + 2|L|\int_0^\ell 6(1-\pi t^2/24)^{n-1}dt.$$

In Figure 3.8 the first integral corresponds to R_1 and R_3 (using polar coordinates) while the second integral corresponds to R_2. For the first integral we have

$$\begin{aligned} 2\pi\int_0^\ell t6(1-\pi t^2/24)^{n-1}dt &= \frac{144}{n}\int_0^\ell \frac{t\,n\pi}{12}(1-\pi t^2/24)^{n-1}dt \\ &= -\frac{144}{n}\left[(1-\pi t^2/24)^n\right]_{t=0}^\ell \\ &\le \frac{144}{n}. \end{aligned}$$

For the second integral we have

$$\begin{aligned} 2|L|\int_0^\ell 6(1-\pi t^2/24)^{n-1}dt &\le 12|L|\int_0^\ell (1-\pi t^2/24)^{n-1}dt \\ &= 12|L|\sqrt{24/\pi}\int_0^{\ell\sqrt{\pi/24}} (1-t^2)^{n-1}dt \\ &\le 24|L|\sqrt{6/\pi}\frac{1}{\sqrt{n}} \end{aligned}$$

since

$$\int_0^1 (1-t^2)^{n-1}dt = \frac{2\cdot 4\cdot 6\cdot\dots\cdot(2n-2)}{3\cdot 5\cdot 7\cdot\dots\cdot(2n-1)} \le \frac{1}{\sqrt{n}} \tag{3.18}$$

and $\ell\sqrt{\pi/24} \le 1$. The latter is the case because L has distance at least 2ℓ to the boundary and therefore C contains a disk with area $4\pi\ell^2$ and C has area 1. For (3.18) we refer to [84] but sketch how it can be derived: the equality in (3.18) can be derived by substituting t by $\sin x$. This gives

$$\int_0^1 (1-t^2)^{n-1}dt = \int_0^{\pi/2} \cos^{2n-1} x dx.$$

For this integral the equality in (3.18) is known (and follows by integration by parts and induction). The inequality in (3.18) follows by induction.

Summarizing, we have in the case of the event $D^c_{w,\ell}$ that the probability for one of the first n points to have a Delaunay edge intersecting L is bounded by

$$24\sqrt{6/\pi}|L|\frac{1}{\sqrt{n}} + 144\frac{1}{n}.$$

For the remaining m points we have by their density bound the same probability with n replaced by m and an additional factor of λ_m. Thus, the expected number of points with Delaunay edges intersecting L is bounded by

$$24\sqrt{6}\pi|L|\frac{n + \lambda_m m}{\sqrt{n}} + 144\frac{n + \lambda_m m}{n}.$$

By Euler's formula the number of intersecting edges is therefore bounded by 3 times this value.

It remains the case that $D_{w,\ell}$ occurs. The total number of edges in the triangulation is bounded by $3(n + m)$ and $\mathrm{P}[D_{w,\ell}] \leq 1/(n + m)$. Therefore this case contributes at most $3(n+m)/(n+m) \leq 3$ to the expected number of intersections. In total we have a bound of

$$72\sqrt{6}\pi|L|\frac{n + \lambda_m m}{\sqrt{n}} + 435\frac{n + \lambda_m m}{n}.$$

□

Thus, the constants c_0, c_1, and c_2 can be bounded independently of C which strengthens the statement of Lemma 3.2. Let us briefly discuss the bound on c_0 because different values for c_0 have been given previously. Devroye, Mücke and Zhu [50] give a value of $c_0 = 3$ in Lemma 3.2 but do not provide a proof. Devroye, Lemaire and Moreau [49] give a value of $c_0 = 10$ for Lemma 3.2 for the special case of the unit square. In the previous lemma we give a value of $c_0 = 4\sqrt{6/\pi} < 5.53$. By bounding $\mathrm{P}[D_{w,\ell}]$ only by $n^{-1/2}$ we can improve this to $c_0 = 4\sqrt{5/\pi} < 5.05$.

Next we generalize Lemma 3.19.

Lemma 3.26. *Let $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ be independently and uniformly distributed points in a convex region C. Let $\pi\colon \{1, \ldots, m\} \to \{1, \ldots, m\}$ be the permutation corresponding to a space-filling curve order of $Y_1, \ldots, Y_m$ for a Hölder-1/2 continuous and bi-measure preserving space-filling curve. For $1 \leq i \leq m - 1$ let I_i denote the number of intersections between the Delaunay triangulation of $X_1, \ldots, X_n, Y_{\pi(1)}, \ldots Y_{\pi(i)}$ and the segment of the space-filling curve tour from $Y_{\pi(i)}$ to $Y_{\pi(i+1)}$. Then*

$$E\left[\sum_{i=1}^{m-1} I_i\right] \in O(m^2/n + n)$$

Proof. Recall that μ_C is the area of C and

$$\mu_s := \mu\left([\psi^*(y), \psi^*(y')] \cap \psi^{-1}(C)\right)$$

for a segment (y, y') of the space-filling curve tour. Without loss of generality let the area of C be 1.

First consider the event D that the space-filling curve tour through $Y_1, \dots, Y_m$ contains a segment s with $\mu_s > 2\frac{\log m}{m-1}\mu_C$. By Lemma 3.22 the probability of this event is bounded from above by $\frac{1}{m}$. In this case we bound for every tour segment, the number of intersections with the Delaunay triangulation by the number of edges in the triangulation which in turn is bounded by $3(n+m)$. Thus, this event contributes less than $3n/m + 3$ to the expected total number of intersections.

Now we consider the case that this event does not occur. In this case we have by Lemma 3.22 that if we condition under $Y_{\pi(i)}$ and $Y_{\pi(i+1)}$ for an $i \in \{1, \dots m-1\}$ then the density function of the distribution of the remaining points of the tour is bounded by $\lambda_m \in O(1 + \log m/m)$.

In the following we will consider for a tour segment $(Y_{\pi(i)}, Y_{\pi(i+1)})$ for $1 \leq i \leq m-1$ the intersection with the Delaunay triangulation of the points $X_1, \dots, X_n, Y_{\pi(1)}, \dots, Y_{\pi(i-1)}$, i.e., the case that even the starting point of the tour segment is not yet inserted. The bound obtained in this way changes when additionally having $Y_{\pi(i)}$ in the triangulation at most by the update cost when inserting $Y_{\pi(i)}$. Thus, summing over all tour segments the additional cost is in $O(m+n) \subset O(m^2/n + n)$.

By Lemma 3.25 the expected number of intersections for a tour segment L with distance at least $c_0\sqrt{\log(n+m)/(n-2)}$ to the boundary is bounded by

$$c_1\frac{n + \lambda_m m}{n} + c_2|L|\frac{n + \lambda_m m}{\sqrt{n}}$$

with c_0, c_1, and c_2 as in Lemma 3.25. Therefore such line segments contribute at most

$$c_1 m\frac{n + \lambda_m m}{n} + c_2\sqrt{m}\frac{n + \lambda_m m}{\sqrt{n}} \in O\left(\sqrt{mn} + \frac{m^2}{n}\right).$$

Next we consider the line segments of the tour near the boundary, i.e., line segment with an endpoint closer than $c_0\sqrt{\log(n+m)/(n-2)}$ to the boundary. The expected number of points on the tour near to the boundary might change by conditioning under the event D but it can increase in this case at most by $1/\mathrm{P}[D] \leq m/(m-1)$. Thus, the expected number of points near the boundary is at most

$$m' := c_0\,|\partial C|\,\frac{m^2}{m-1}\,\sqrt{\log(n+m)/(n-2)}.$$

Choosing for $n > 2$

$$\begin{aligned} w &:= (\log(n+m)/(n-2))^{2/3} \quad \text{and} \\ \ell &:= 6\sqrt{2}(\log(n+m)/(n-2))^{1/3}. \end{aligned}$$

gives using Lemma 3.23 (as in the proof of Lemma 3.19) as bound on the number of intersections near the boundary

$$
\begin{aligned}
&3(n + \lambda_m m)(|\partial C| 2m' w + \pi \ell^2 2m' + 2\ell c_T \sqrt{3m'}) \\
&\quad \in \ O\left((n+m)(m(\log(n+m)/n)^{7/6} + \sqrt{m}(\log(n+m)/n)^{5/6}) \right) \\
&\quad \subset \ O\left(\sqrt{mn} + \frac{m^2}{n} \right)
\end{aligned}
$$

with probability at least $1 - 1/(n+m)^2$. The remaining case contributes at most $O(1)$ to the expected number of intersections. As in the proof of Lemma 3.19 this can be seen by taking the worst-case bound of $3(n+m)$ intersections per tour segment. □

Theorem 3.27. *Using a biased randomized insertion order and, in each round, walking along a Hölder-$1/2$, bi-measure preserving space-filling curve, the incremental construction algorithm runs in linear expected time for points distributed independently and uniformly at random in a bounded convex region, assuming linear time bucketing for the space-filling curve computation.*

Proof. By Corollary 2.10, a space-filling curve order can be computed in linear expected time and by Proposition 3.7 the expected update cost is linear. We therefore only need to bound the expected time for point location.

We analyze only the last round. The overall linearity then follows from the linearity of expectation. For analyzing the last round we use Lemma 3.26. It remains to see for N points that the expected value of m^2/n is linear in N where n comes from a binomial distribution B with N tries and probability $1/\alpha$ and m equals $N - n$. The expected value of m^2/n is

$$
\mathrm{E}\left[\frac{(N-B)^2}{B}\right] = N^2\,\mathrm{E}\left[\frac{1}{B}\right] - 2N - \mathrm{E}\left[B\right].
$$

By Jensen's inequality and by inserting $\mathrm{E}[B] = N/\alpha$ we obtain

$$
\mathrm{E}\left[\frac{(N-B)^2}{B}\right] \ \le \ \frac{N^2}{N/\alpha} - 2N - N/\alpha \ = \ (\alpha + 1/\alpha - 2)N.
$$

Finally, it is necessary to take into account that the starting point of each line segment walk is a vertex of the Delaunay triangulation and we therefore need to bound its degree to find the first triangle stabbed by the line segment. This can be bounded by the update cost which is expected to be linear [3]. □

Conclusion

The two main results of this chapter are a generalized analysis of incremental constructions con BRIO and the linear expected running time of incremental constructions con BRIO when used with a space-filling curve order on independent uniformly distributed points in a convex region in the plane.

The first result describes the worst-case behaviour of incremental constructions con BRIO in terms of the degree of the configuration space and the sampling parameter α. The result suggests to use a small sampling parameter α, i.e., a large sampling ratio. We will see in the next chapter that this is not necessarily the case because further trade-offs have to be taken into account.

The second result describes the average-case behaviour of the algorithm on uniformly distributed points in the plane. In the next chapter we will analyze the running time of the algorithm on uniformly distributed points in higher dimension. Nonetheless, the results of the next chapter do not fully generalize the result obtained here. First, the analysis here holds for arbitrary convex regions while the analysis of the next chapter applied to uniformly distributed points holds for points in a d-cube. Second and more important, the analysis of this chapter also holds for other traversals of the points than space-filling curve orders. An open problem is to obtain a similar result in higher dimensions.