



HAL
open science

Efficiently navigating a random Delaunay triangulation

Nicolas Broutin, Olivier Devillers, Ross Hemsley

► **To cite this version:**

Nicolas Broutin, Olivier Devillers, Ross Hemsley. Efficiently navigating a random Delaunay triangulation. *Random Structures and Algorithms*, 2016, 49 (1), pp.95–136. 10.1002/rsa.20630 . hal-00940743v3

HAL Id: hal-00940743

<https://inria.hal.science/hal-00940743v3>

Submitted on 11 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Efficiently Navigating a Random Delaunay Triangulation ^{*}

Nicolas Broutin [†] Olivier Devillers [‡] Ross Hemsley [§]

April 11, 2016

Abstract

Planar graph navigation is an important problem with significant implications to both point location in geometric data structures and routing in networks. However, whilst a number of algorithms and existence proofs have been proposed, very little analysis is available for the properties of the paths generated and the computational resources required to generate them under a random distribution hypothesis for the input. In this paper we analyse a new deterministic planar navigation algorithm with constant spanning ratio (w.r.t the Euclidean distance) which follows vertex adjacencies in the Delaunay triangulation. We call this strategy *cone walk*. We prove that given n uniform points in a smooth convex domain of unit area, and for any start point z and query point q ; cone walk applied to z and q will access at most $O(|zq|\sqrt{n} + \log^7 n)$ sites with complexity $O(|zq|\sqrt{n} \log \log n + \log^7 n)$ with probability tending to 1 as n goes to infinity. We additionally show that in this model, cone walk is $(\log^{3+\xi} n)$ -memoryless with high probability for any pair of start and query point in the domain, for any positive ξ . We take special care throughout to ensure our bounds are valid even when the query points are arbitrarily close to the border.

1 Introduction

Given a planar embedding of a graph $G = (V, E)$, a source node $z \in V$ and a destination point $q \in \mathbb{R}^2$, we consider the *planar graph navigation* problem of finding a route in G from z to the nearest neighbour of q in V . In particular, we assume that any vertex $v \in V$ may access its coordinates in \mathbb{R}^2 with a constant time query. The importance of this problem is two-fold. On the one hand, finding a short path between two nodes in a network is currently a very active area of research in the context of routing in networks [1, 22, 27]. On the other, the problem of locating a face containing a point in a convex subdivision (point location) is an important sub-routine in many algorithms manipulating geometric data structures [11, 12, 18, 23]. A number of algorithms have been proposed within each of these fields, many of which are in fact equivalent. It seems the majority of the literature in these areas is concerned with the existence

^{*}The work in this paper has been partially supported by ANR blanc PRESAGE (ANR-11-BS02-003)

[†]Projet RAP, INRIA Paris - Rocquencourt

[‡]Projet Geometrica, INRIA Sophia Antipolis - Méditerranée

[§]Projet Geometrica, INRIA Sophia Antipolis - Méditerranée

of algorithms which always succeed under different types of constraints, such as the *spanning ratio* or the *competitiveness* of the algorithm, or the class of network. Apart from worst-case bounds, very little is known concerning the properties of the path lengths and running times for these algorithms under random distribution hypotheses for the input vertices. In this paper we aim to bridge the gap between these two fields by giving and analysing an algorithm that is provably efficient within both of these contexts when the underlying graph is the Delaunay triangulation.

1.1 Definitions

In the following, we define the *spanning ratio* of an algorithm to be the worst case ratio between the length of the path generated by the algorithm and the Euclidean distance between the source and the destination and the *competitiveness* the worst case ratio between the lengths of the paths produced by the studied algorithm and an optimal offline algorithm. Thus the spanning ratio is always larger than the competitiveness and a good spanning ratio imply a good competitiveness. They both may depend on the class of graphs one allows, but not the pair of source and destination. Let $\mathbf{N}_d(v)$ denote the set of neighbours of v within d hops of v . We shall sometimes refer to the d 'th neighbourhood of a set X , to denote the set of all sites that can be accessed from a site in X with fewer than d hops. We call an algorithm *c-memoryless* if at each step in the navigation, it only has access to the destination q , the current vertex v and $\mathbf{N}_c(v)$. Some authors use the term *online* to refer to an algorithm that only has access to q , the current vertex v , $\mathbf{N}_1(v)$ and $O(1)$ words of memory which may be used to store information about the history of the navigation. Finally, an algorithm may be either deterministic or randomised. We define a randomised algorithm to be an algorithm that has access to a random oracle at each step.

1.2 Previous results

Graph Navigation for Point location. The problem of point location is most often studied in the context of triangulations and the algorithms are referred to as *walking algorithms* [12]. A walking algorithm may work by following edges or by following incidences between neighbouring faces, which is equivalent to a navigation in the dual graph. There are three main algorithms that have received attention in the literature: *straight walk*, which is a walk that visits all triangles crossed by the line segment zq ; *greedy vertex walk*, which always chooses the vertex in $\mathbf{N}_1(v)$ which is closest to q and *visibility walk* which walks to an adjacent triangle if and only if it shares the same half-space as q relative to the shared edge. It is known that these algorithms always terminate if the underlying triangulation is Delaunay [12].

The aim is generally to analyse the expected number of steps that the algorithm requires to reach the destination under a given distribution hypothesis. Such an analysis has only been provided by Devroye et al. [14] who succeeded in showing that straight walk reaches the destination after $O(\|zq\|\sqrt{n})$ steps in expectation, for n random points in the unit square (here and from now on we shall use $\|\cdot\|$ to denote the Euclidean distance).¹ The analysis in this case is

¹ Zhu also provides a tentative $O(\sqrt{n \log n})$ bound for visibility walk [28]. This is a proof by induction for “a random edge at distance d ”. It considers the next edge in the walk and applies an induction hypothesis to try and bound the progress. Unfortunately, the new edge cannot be considered as random: each edge that is visited has been

facilitated since it is possible to compute the probability that a triangle is part of the walk without looking at the other vertices. Straight walk is online, but not memoryless since at every step the algorithm must know the location of the source point, z . It is also rarely used in practice since it is usually outperformed empirically by one of the remaining two algorithms, *visibility walk* or *greedy vertex walk*, which are both 1-memoryless [12]. The complex dependence between the steps of the algorithm in these cases makes the analysis difficult, and it remains an important open question to provide an analysis for either of these two algorithms.

Graph Navigation for Routing. In the context of packet routing in a network, each vertex represents a node which knows its approximate location and can communicate with a selected set of neighbouring nodes. One example is in wireless networks where a node communicates with all devices within its communication range. In such cases, it is often convenient for the nodes to agree on a communication protocol such that the graph of directly communicating nodes is planar, since this can make routing more efficient. Triangulations have been used in this context due to their ability to act as spanners (the length of shortest paths in the graph, seen as curves in \mathbb{R}^2 , do not exceed the Euclidean distance by more than a constant factor), and methods exist to locally construct the full Delaunay triangulation, given some conditions on the point distribution [17, 19, 24].

Commonly referenced algorithms in this field are: *greedy routing*, which is the same algorithm as *greedy vertex walk*, given in the context of point location; *compass routing* which is similar to greedy, except that instead of choosing the point in $\mathbf{N}_1(v)$ minimising the distance to q , it chooses the point in $x \in \mathbf{N}_1(v)$ minimising the angle $\angle q, v, x$, and also *face routing* which is a generalisation of straight walk that can be applied to any planar graph. In this context, overall computation time is usually considered less important than trying to construct algorithms that find short paths in a given network topology under certain memory constraints. We give a brief overview of results relating to this work.

Bose et al. [8] demonstrated that it is not possible to construct a deterministic memoryless algorithm that finds a path with constant competitiveness in an arbitrary triangulation. They also demonstrated by counter example that neither greedy routing, nor compass routing is $O(1)$ -competitive on the Delaunay triangulation [6]. Bose and Morin [7] went on to show that there does, however, exist an online c -competitive algorithm that works on any graph satisfying a property they refer to as the ‘diamond property’, which is satisfied by Delaunay triangulations. They show this by providing an algorithm which is essentially a modified version of the straight walk. Bose and Morin also show that there is no algorithm that is competitive for the Delaunay triangulation under the link length (the link length is the number of edges visited by the algorithm) [8].

In terms of time analysis, it appears the only relevant results are those by Devroye et al. [14], where their results correspond with that of the straight walk, and those by Chen et al. [10], who show that no routing algorithm is asymptotically better than a random walk when the underlying graph is an arbitrary convex subdivision.

chosen by the algorithm and the edges do not all have the same probability to be chosen at each step in the walk. Restarting the walk from a given edge is not possible either (as done in [13]), since the knowledge that an edge is a Delaunay edge influences the local point distribution.

Navigation in the Plane. We briefly remark that for the related problem of navigation in the plane, several probabilistic results exist; for example [3] and [4]. In this context, the input is a set of vertices in the plane along with an oracle that can compute the next step given the current step and the destination in $O(1)$ time. Although the steps are also dependent in these cases, the case of Delaunay triangulations we treat here is more delicate because of the geometry of the region of dependence implied by the Delaunay property.

1.3 Contributions

In this paper we give a new deterministic planar graph navigation algorithm which we call *cone walk* that succeeds on any Delaunay triangulation and produces a path which is 3.7-spanning ratio. We briefly underline the fact that our algorithm has been designed for theoretical demonstration, and we do not claim that it would be faster in a practical sense than, for example, *greedy routing* or *face routing*. On the other hand, direct comparisons would perhaps be unfair, since *greedy routing* is not $O(1)$ -competitive on the Delaunay triangulation [6] whereas we prove that cone walk is; and *face routing* is not memoryless in any sense, whilst cone walk is localised in the sense given by Theorem 1. In the theorems that follow, we characterise the asymptotic properties of the cone walk algorithm applied to a random input.

Let \mathcal{D} be a smooth convex domain of the plane with area 1, and write $\mathcal{D}_n = \sqrt{n}\mathcal{D}$ for its scaling to area n . For $x, y \in \mathcal{D}$, let $\|xy\|$ denote the Euclidean distance between x and y . Under the hypothesis that the input is the Delaunay triangulation of n points uniformly distributed in a convex domain of unit area, we prove that, for any $\varepsilon > 0$, our algorithm is $O(\log^{1+\varepsilon} n)$ -memoryless with probability tending to one. In the case of cone walk, this is equivalent to bounding the number of neighbourhoods that might be accessed during a step, which we deal with in the following theorem.

Theorem 1. *Let $\mathbf{X}_n := \{X_1, X_2, \dots, X_n\}$ be a collection of n independent uniformly random points in \mathcal{D}_n . For $z \in \mathbf{X}_n$ and $q \in \mathcal{D}_n$, let $M(z, q)$ be the maximum number of neighbourhoods needed to compute every step of the walk. Then, for every $\varepsilon > 0$,*

$$\mathbb{P}\left(\exists z \in \mathbf{X}_n, q \in \mathcal{D}_n : M(z, q) > \log^{3+\varepsilon} n\right) \leq \frac{1}{n}.$$

In particular, as $n \rightarrow \infty$, $\mathbb{E}[\sup_{z \in \mathbf{X}_n, q \in \mathcal{D}_n} M(z, q)] = O(\log^{3+\varepsilon} n)$, for every $\varepsilon > 0$.

Also with probability close to one, we show that the path length, the number of edges and the number of vertices accessed are $O(\|zq\| + \log^6 n)$ for any pair of points in the domain. We formalise these properties in the following theorem.

Theorem 2. *Let $\mathbf{X}_n := \{X_1, X_2, \dots, X_n\}$ be a collection of n independent uniformly random points in \mathcal{D}_n . Let $\Gamma(z, q)$ denote either the Euclidean length of the path generated by the cone walk from $z \in \mathbf{X}_n$ to $q \in \mathcal{D}_n$, its number of edges, or the number of vertices accessed by the algorithm when generating it. Then there exist constants $C_{\Gamma, \mathcal{D}}$ depending only on Γ and on the shape of \mathcal{D} such that, for all n large enough,*

$$\mathbb{P}\left(\exists z \in \mathbf{X}_n, q \in \mathcal{D}_n : \Gamma(z, q) > C_{\Gamma, \mathcal{D}} \cdot \|zq\| + 4(1 + \sqrt{\|zq\|}) \log^6 n\right) \leq \frac{1}{n}.$$

In particular, as $n \rightarrow \infty$,

$$\mathbb{E} \left[\sup_{z \in \mathbf{X}_n, q \in \mathcal{D}_n} \Gamma(z, q) \right] = O(\sqrt{n}).$$

Finally, we bound the computational complexity of the algorithm, $T(z, q)$.

Theorem 3. *Let $\mathbf{X}_n := \{X_1, X_2, \dots, X_n\}$ be a collection of n independent uniformly random points in \mathcal{D}_n . Then in the RAM model of computation, there exists a constant C depending only on the shape of \mathcal{D} and the particular implementation of the algorithm such that for all n large enough,*

$$\mathbb{P} \left(\exists z \in \mathbf{X}_n, q \in \mathcal{D}_n : T(z, q) > C \cdot \|zq\| \log \log n + (1 + \sqrt{\|zq\|}) \log^6 n \right) \leq \frac{1}{n}.$$

In particular, as $n \rightarrow \infty$,

$$\mathbb{E} \left[\sup_{z \in \mathbf{X}_n, q \in \mathcal{D}_n} T(z, q) \right] = O(\sqrt{n} \log \log n).$$

Remark 1. The choice of the initial vertex is never discussed. However, previous results show that choosing this point carefully can result in an expected asymptotic speed up for any graph navigation algorithm [23].

1.4 Layout of the paper

In Section 2, we give a precise definition of the cone walk algorithm and prove some important geometric properties. In Section 3, we begin the analysis for the cone walk algorithm applied to a homogeneous planar Poisson process. To avoid problems when the walk goes close to the boundary, we provide an initial analysis which assumes that the points are sampled from a disc with the query point at its centre. This analysis is then extended to arbitrary query points in the disc and also to other convex domains in Section 4. In Section 5, we prove estimates about an auxiliary line arrangement which are crucial to proving the worst-case probabilistic bounds in Theorems 1 and 3. Finally, we compare our findings with computer simulations in Section 6.

2 Algorithm and geometric properties

We consider the finite set of sites in general position (so that no three points of the domain are co-linear, and no four points are co-circular), $\mathbf{X} \subset \mathbb{R}^2$ contained within a compact convex domain $\mathcal{D} \subset \mathbb{R}^2$. Let $\text{DT}(\mathbf{X})$ be the Delaunay triangulation of \mathbf{X} , which is the graph in which three sites $x, y, z \in \mathbf{X}$ form a triangle if and only if the disc with x, y and z on its boundary does not contain any site in \mathbf{X} . Given two points $z, q \in \mathbb{R}^2$ and a number $r \in \mathbb{R}$ we define $\text{Disc}(z, q, r)$ to be the closed disc whose diameter spans z and the point at a distance $2r$ from z on the ray zq . Finally, we define $\text{Cone}(z, q, r)$ to be the sub-region of $\text{Disc}(z, q, r)$ contained within a closed cone of apex z , axis zq and half angle $\frac{\pi}{8}$ (see Figure 1).

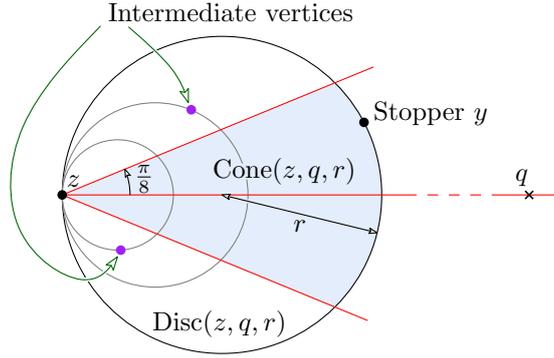


Figure 1: Choosing the next vertex.

2.1 The cone walk algorithm

Given a site $z \in \mathbf{X}$ and a destination point $q \in \mathcal{D}$, we define one *step* of the cone walk algorithm by growing the region $\text{Cone}(z, q, r)$ anchored at z from $r = 0$ until the first point $y \in \mathbf{X}$ is found such that the region is non-empty. Once y has been determined, we refer to it as the *stopper*. We call the region $\text{Cone}(z, q, r)$ for the given r a *search cone*, and we call the associated disc $\text{Disc}(z, q, r)$ the *search disc* (see Figure 1). The point y is then selected as the anchor of a new search cone $\text{Cone}(y, q, \cdot)$ and the next step of the walk begins. See Figure 2 for an example run of the algorithm.

To find the stopper using only neighbour incidences in the Delaunay triangulation, we need only access vertices in a well-defined local neighbourhood of the search disc. Define the points $\mathbf{X} \cap \text{Disc}(z, q, r) \setminus \{z, y\}$ to be the *intermediate vertices*. The algorithm finds the stopper at each step by gradually growing a disc anchored at z in the direction of the destination, adding the neighbours of all vertices in \mathbf{X} intersected along the way. This is achieved in practice by maintaining a series of candidate vertices initialised to the neighbours of z and selecting amongst them the vertex defining the smallest search disc at each iteration. Each time we find a new vertex intersecting this disc, we check to see if it is contained within $\text{Cone}(z, q, \infty)$. If it is, this point is the next stopper and this step is finished. Otherwise the point must be an intermediate vertex and we add its neighbours to the list of candidate vertices. This procedure works because the intermediate vertex defining the next largest disc is always a neighbour of one of the intermediate vertices that we have already visited during the current step (see Lemma 5).

We terminate the algorithm when the destination q is contained within the current search disc for a given step. At this point we know that one of the points contained within $\text{Disc}(z, q, r)$ is a Delaunay neighbour of q in $\text{DT}(\mathbf{X} \cup \{q\})$. We can further compute the triangle of $\text{DT}(\mathbf{X})$ containing the query point q (point location) or find the nearest neighbour of q in $\text{DT}(\mathbf{X})$ by simulating the insertion of the point q into $\text{DT}(\mathbf{X})$ and performing an exhaustive search on the neighbours of q in $\text{DT}(\mathbf{X} \cup \{q\})$.

We will sometimes distinguish between the *visited* vertices, which we take to be the set of all sites contained within the search discs for every step and the *accessed* vertices, which we define to be the set of all vertices accessed by the cone-walk algorithm. Thus the accessed vertices are

the visited vertices along with their 1-hop neighbourhood.

The pseudo-code below gives a detailed algorithmic description of the CONE-WALK algorithm. We take as input some $z \in \mathbf{X}$, $q \in \mathcal{D}$ and return a Delaunay neighbour of q in $\text{DT}(\mathbf{X} \cup \{q\})$. Recalling that $\mathbf{N}_1(v)$ refers to the Delaunay neighbours of $v \in \text{DT}(\mathbf{X})$ and additionally defining $\text{NEXT-VERTEX}(S, z, q)$ to be the procedure that returns the vertex in S with the smallest r such that $\text{Disc}(z, q, r)$ touches a vertex in S and $\text{IN-CONE}(z, q, y)$ to be `true` when $y \in \text{Cone}(z, q, \infty)$.

```

CONE-WALK( $z, q$ )
1   $Substeps = \{z\}$ 
2   $Candidates = \mathbf{N}_1(z)$ 
3  while true
4       $y = \text{NEXT-VERTEX}(Candidates \cup \{q\}, z, q)$ 
5      if  $\text{IN-CONE}(z, q, y)$ 
6          if  $y = q$ 
7              // Destination reached.
8              return  $\text{NEXT-VERTEX}(Substeps, q, z)$ 
9              //  $y$  is a stopper
10              $z = y$ 
11              $Substeps = \{z\}$ 
12              $Candidates = \mathbf{N}_1(z)$ 
13         else
14             //  $y$  is an intermediate vertex.
15              $Substeps = Substeps \cup \{y\}$ 
16              $Candidates = Candidates \cup \mathbf{N}_1(y) \setminus Substeps$ 

```

2.2 Path Generation

We note that the order in which the vertices are discovered during the walk does not necessarily define a path in $\text{DT}(\mathbf{X})$. If we only wish to find a point of the triangulation that is close to the destination (for example, in point location), this is not a problem. However, in the case of routing, a path in the triangulation is required to provide a route for data packets. To this end, we provide two options that we shall refer to as `SIMPLE-PATH` and `COMPETITIVE-PATH`. `SIMPLE-PATH` is a simple heuristic that can quickly generate a path that is provably short on average. `SIMPLE-PATH` may have an unbounded spanning ratio in some pathological situations where the path makes zig-zag, it has only a good expected spanning-ratio. `COMPETITIVE-PATH` is slightly more complex from an implementation point of view, however we show that for any possible input the algorithm will always generate a path of constant spanning ratio whilst still maintaining the same asymptotic behaviour under the point distribution hypotheses explored in Section 3.

`SIMPLE-PATH` A simple way to generate a valid path is to keep a predecessor table for each vertex. We start with an empty table at the beginning of each step, and then every time we access a new vertex, we store it in the table along with the vertex that we accessed it from. To trace a path back, we simply follow the predecessors.

COMPETITIVE-PATH Let Z_i for $i > 0$, be the i th stopper in the walk thus, Z_i is the stopper found at step i) and $Z_0 := z$. For a path to have a good spanning ratio, it should at least have a good spanning ratio *locally*: for each step, there should be a bound on the length of the path generated between Z_i and Z_{i+1} , which does not depend on the points in the search disc. To construct a path verifying this property, we use the fact that the *stretch factor* of the Delaunay triangulation is bounded above by a constant, λ . This means that for any two sites x, y , there exists a path from x to y in the Delaunay triangulation for which the sum of the lengths of the edges is at most $\lambda\|xy\|$. Currently the literature gives us that the stretch factor is in $[1.5932, 1.998]$ [25, 26]. Clearly this implies that there exists a path between Z_i and Z_{i+1} with total length at most $\lambda\|Z_i Z_{i+1}\|$. Moreover, for $\lambda = 1.998$, Xia proved that there exists such a path using only vertices of the Delaunay edges crossing the line segment xy , and such edges have at least one vertex in the intermediate vertices between Z_i and $Z_i + 1$. We use Dijkstra's algorithm to find the shortest path between Z_i and Z_{i+1} which uses only intermediate vertices between Z_i and Z_{i+1} and their neighbors. The resulting path implicitly has stretch bounded by λ . We show in Lemma 4 that this algorithm results in a bound for the spanning ratio for the full path.

Lemma 4. *CONE-WALK has 3.7-spanning ratio when the COMPETITIVE-PATH algorithm is used to generate the path in $\text{DT}(\mathbf{X})$*

Proof. Let Z_i, Z_{i+1} be the stoppers of two consecutive steps defined by the algorithm. The stretch factor bound guarantees that the path generated between Z_i and Z_{i+1} has length bounded by $\lambda\|Z_i Z_{i+1}\|$, meaning that the longest path can have stretch at most $\lambda \sum_{i=0}^{\tau-1} \|Z_i Z_{i+1}\| / \|zq\|$ where τ is the number of steps in the walk. We bound this sum by observing that $\|Z_i Z_{i+1}\| \leq 2 \cos \frac{\pi}{8} \cdot (\|Z_i q\| - \|Z_{i+1} q\|)$, which follows from Figure 8. Finally, no path defined by the algorithm can be longer than

$$\lambda \sum_{i=0}^{\tau-1} \|Z_i Z_{i+1}\| \leq 2\lambda \cos \frac{\pi}{8} \sum_{i=0}^{\tau-1} (\|Z_i q\| - \|Z_{i+1} q\|) \leq 2\lambda \cos \frac{\pi}{8} \cdot \|zq\|$$

Thus the path has a spanning ratio of c for $c := 2\lambda \cos \frac{\pi}{8} \leq 4 \cos \frac{\pi}{8} \leq 3.7$. \square

2.3 Complexity

In this section we give deterministic bounds on the number of operations required to compute $\text{CONE-WALK}(z, q)$ within the RAM model of computation. In this model, accessing, comparing and performing arithmetic on points is treated as atomic. We will use these deterministic bounds to extract probabilistic bounds under certain distribution assumptions in Section 3.4. For now, we focus on a single step of the walk starting from $y \in \mathbf{X}$, and resulting in a disc with radius r . Let k be the number of points intersecting the disc $\text{Disc}(y, q, r)$ and m be the number of edges in $\text{DT}(\mathbf{X})$ intersecting $\partial \text{Disc}(y, q, r)$ (where we use the notation ∂A to denote the boundary of A).

We note that every intermediate vertex will add its neighbours to the list of *Candidates* when visited. Each of these insertions can be associated with a single edge of $\text{DT}(X)$ intersecting $\text{Disc}(y, q, r)$ (with multiplicity two for each 'internal' edge, since they are accessed from

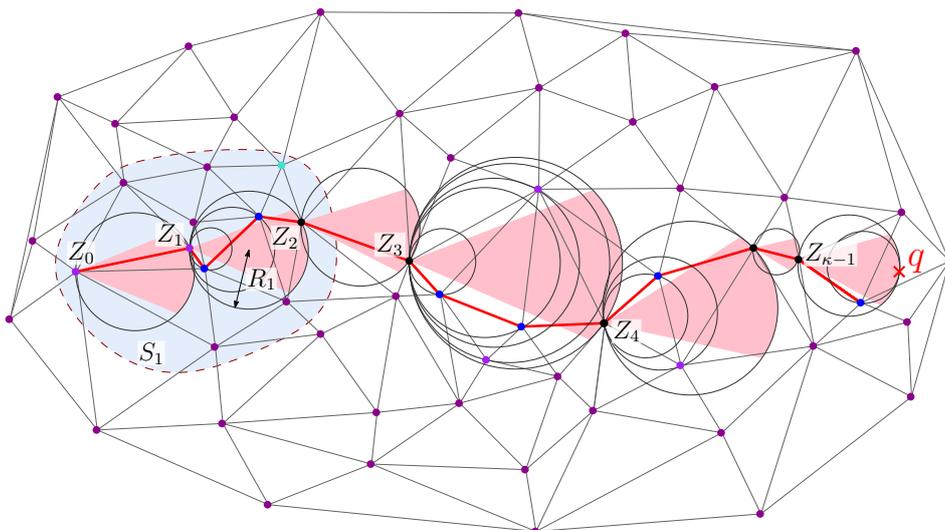


Figure 2: An example of cone walk. The points $(Z_i)_{i>0}$ are the *stoppers* (or sometimes, the *steps*) and the points within each circle are the *intermediate vertices*. The shaded conic regions are the *cones*, and the set of outer circles for each of the steps in the diagram is referred to as the *discs*.

both sides). By the Euler relation, the total number of such insertions for one step is thus at most $3(m + 2k)$. In addition, we observe that when moving from one intermediate vertex to the next, a search in the list of *Candidates* is required. A simple linear search requires $O(m + k)$ operations for each intermediate vertex. Combining this with the above, we achieve a bound of $O(k(m + k))$ operations for one step. This bound may be improved by replacing *Candidates* with a priority queue keyed on the associated search-disc radius of each candidate, which yields a simple improvement to $O(k \log(m + k))$.

For the path generation algorithms, we observe that SIMPLE-PATH only requires a constant amount of processing per vertex accessed to generate the predecessor table and $O(k)$ time to output the path at the end of each step, so the asymptotic running time is not affected by its inclusion. COMPETITIVE-PATH is slightly more complicated since it uses Dijkstra's algorithm on the k points in the search disc and their neighbors. Let k' be the number of neighbors, the complexity is $O((k + k') \log(k + k'))$.

2.4 Geometric properties

We now prove a series of geometric lemmata giving properties of steps in the walk. We begin with a small lemma that will guarantee that we never get 'stuck' when performing a search for the next step, thus demonstrating correctness of the algorithm. The following two 'overlapping' lemmata allow us to establish which regions may be considered independent in a probabilistic sense and will be important in Section 3. Finally we provide a 'stability' result, which will help us to bound the region in which a destination point may be moved without changing the

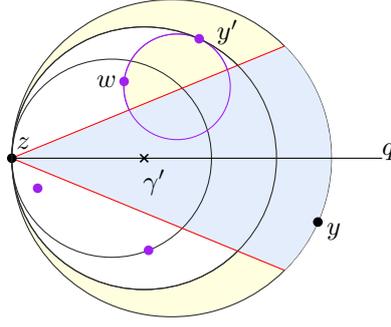


Figure 3: We observe that y' has always a Delaunay neighbour in $\text{Disc}(z, q, r)$, where r is the radius ensuring $y \in \partial\text{Disc}(z, q, r)$.

sequence of steps taken by the algorithm. This will be important when we enumerate the number of different walks possible for a given set of input points.

2.4.1 Finding a Delaunay path within the discs

Lemma 5 (Path finding lemma). *Let $q \in \mathcal{D}$, $z \in \mathbf{X}$ and $y' \in \mathbf{X}$ with associated disc $\text{Disc}(z, q, r')$. Suppose there exists an $r > 0$ such that $(\text{Disc}(z, q, r') \setminus \text{Disc}(z, q, r)) \cap \mathbf{X} = \{y'\}$. Then there exists a point in $\text{Disc}(z, q, r)$ that is a Delaunay neighbour of y' .*

Proof. Let γ' be the centre of $\text{Disc}(z, q, r')$. We grow $\text{Disc}(y', \gamma', \rho) \subset \text{Disc}(z, q, r')$ until we hit a point w in \mathbf{X} . The point w is always contained within $\text{Disc}(z, q, r)$ because z is on the border of $\text{Disc}(z, q, r)$. Since the interior of the $\text{Disc}(y', \gamma', \rho)$ is empty, w is a Delaunay neighbour of y' . See Figure 3. \square

Corollary 6. *Let $q \in \mathcal{D}$, $z \in \mathbf{X}$ with $y \in \mathbf{X}$ its associated stopper satisfying $y \in \partial\text{Cone}(z, q, r)$. Then there is a path of edges of $\text{DT}(\mathbf{X})$ from z to y contained within $\text{Disc}(z, q, r)$.*

2.4.2 Independence of the search cones

When growing a new search cone, it is important to observe that it does not overlap any of the previous search cones, except at the very end of the walk. This is formalised by the following lemma.

Lemma 7 (Non-overlapping lemma). *Let z and y be two points of \mathbf{X} and $r > 0$ such that $\text{Cone}(z, q, r)$ has y on its boundary. If $\|zq\| > (2 + \sqrt{2})r$ then $\text{Disc}(z, q, r)$ does not intersect the search cone $\text{Cone}(y, q, \infty)$ issued from y nor any other search cone for any subsequent step of the walk.*

Proof. Assume without loss of generality that y lies to the left of line zq and consider the construction given in Figure 4. Let β denote the angle between the tangent to $\text{Disc}(z, q, r)$ at y and the ray bordering $\text{Cone}(y, q, \infty)$. $\text{Cone}(y, q, \infty)$ and $\text{Disc}(z, q, r)$ do not intersect provided that

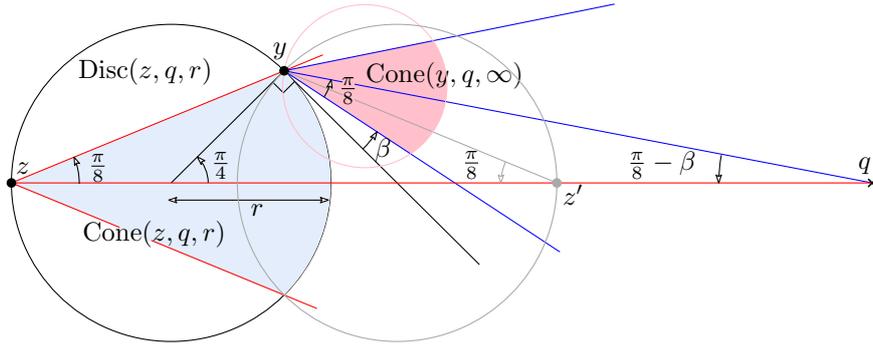


Figure 4: For the proof of Lemma 7.

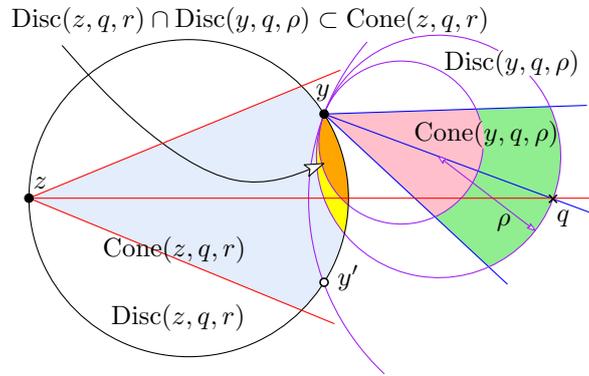


Figure 5: For the proof of Lemma 8.

$\beta \geq 0$. Placing y at the corner of $\text{Cone}(z, q, r)$ maximizes β , in which case we have $\beta > 0$ if and only if q is to the right of z' , the point symmetrical to z with respect to the line through y perpendicular to zq . Elementary computations then yield the result. Since the whole sequence of search cones following the one issued from y remains in $\text{Cone}(y, q, \infty)$, $\text{Disc}(z, q, r)$ does not intersect any of these search cones, and the result follows. \square

2.4.3 Independence of the search discs

When growing the search disc region, the new search disc may overlap previous search discs but only in their cone parts. This is formalised by the following lemma:

Lemma 8 (Overlapping lemma). *Let z and y be two points such that $\text{Cone}(z, q, r)$ has y on its boundary. Then if the search disc $\text{Disc}(y, q, \rho)$ issued from y does not contain q , it does not intersect $\text{Disc}(z, q, r) \setminus \text{Cone}(z, q, r)$.*

Proof. By symmetry we observe that $\text{Disc}(y, q, \rho)$ only intersects the point y' , the point y reflected through the line zq , when the centre of $\text{Disc}(y, q, \rho)$ coincides with q (Figure 5). Since

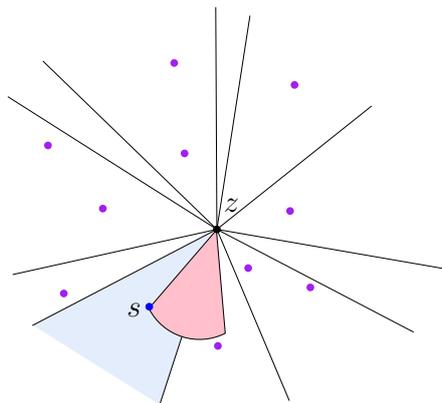


Figure 6: For the proof of Lemma 9. For a given step z , moving the destination in the shaded sector will always result in the same stopper, s , being chosen for the next step.

the algorithm terminates as soon as the current search disc touches q , q is never contained within $\text{Disc}(z, q, \rho)$ and thus this can never happen. \square

2.4.4 Stability of the walk

In the following lemma we are interested in the stability of the sequence of steps to reach q .

Lemma 9 (Invariance lemma). *For an n -set $\mathbf{X} \subseteq \mathbb{R}^2$, there exists an arrangement of half-lines $\Xi = \Xi(\mathbf{X})$ such that the associated subdivision of the plane has fewer than $2n^4$ cells, and such that the sequence of steps used by the cone walk algorithm from any vertex of \mathbf{X} does not change when the aim q moves in a connected component of $\mathbb{R}^2 \setminus \Xi$.*

Proof. Take a point $z \in \mathbf{X}$ and consider \mathcal{S}_z , the set of all possible stoppers defined by $\text{Cone}(z, q, r)$ for some $q \in \mathcal{D}$ and $r > 0$. Each $s \in \mathcal{S}_z$ defines a unique sector about z such that moving a point in the given sector does not change the stopper (see Figure 6). We then create an arrangement by adding a ray on the border of every sector for each point $z \in \mathbf{X}$. The resulting arrangement has the property that moving the destination point q within one of the cells of the arrangement does not change the stopper of any step for any possible walk. Clearly, $|\mathcal{S}_z| \leq n - 1$ for all $z \in \mathbf{X}$, and each sector is bounded by at most two rays, thus there are at most $2n(n - 1)$ rays in the arrangement. Since an arrangement of m lines has at most $\frac{m^2 + m + 2}{2}$ cells the result follows [see, e.g., 20, p. 127]. \square

3 Cone walk on Poisson Delaunay in a disc

Our aim in this section is to prove the main elements towards Theorem 3, which we go on to complete in Section 4. Our ultimate goal is to prove bounds on the behaviour of the cone walk for the *worst* possible pair of starting point and query when the input sites are generated by a

homogeneous Poisson process in a compact convex domain. Achieving this requires first strong bounds on the probability that the walk behaves badly for a fixed start point and query. One then proves the worst-case bounds by showing that to control every possible run of the algorithm, it suffices to bound the behaviour of the walk for enough pairs of starting points and query; this relies crucially on the arrangement of Lemma 9. The tail bounds required in the second stage of the proof may not be obtained from Markov or Chebyshev’s inequalities together with mean or variance estimates only, and we thus need to resort to stronger tools.

Our techniques rely on *concentration inequalities* [9, 15, 16, 21]. Most of the bounds we obtain (for the number of steps κ and the number of visited sites) follow from a representation as a sum of random variables in which the increments can be made independent by a simple and natural conditioning. The bounds on the complexity of the algorithm CONE-WALK are slightly trickier to derive because there is no way to make the increments independent.

For the sake of presentation, we introduce two simplifications which we remove in Section 4. First, we start by studying the walk in the disc \mathcal{D}_n of area n where the query is at the centre. These choices for \mathcal{D}_n and q ensure that for any $z \in \mathcal{D}_n$ and any $r \leq \sqrt{n/\pi}$, we have $\text{Disc}(z, q, r) \subset \mathcal{D}_n$. Note that since the distance to the aim is decreasing, the disc is precisely the *effective* domain where the walk from z and aiming at q takes place.

Then, we introduce independence between the different regions of the domain by replacing the collection of independent points \mathbf{X}_n by a (homogeneous) Poisson point process Φ and consider $\text{DT}(\Phi)$. Recall that a Poisson point process of intensity 1 is a random collection of points $\Phi \subset \mathcal{D}_n$ such that with probability one, all the points are distinct, for any two Borel sets $R, S \subseteq \mathcal{D}_n$, the number of points $|\Phi \cap R|$ is distributed like a Poisson random variable whose mean is the area $\mathcal{A}(R)$ of R , and if $R \cap S = \emptyset$ then $|\Phi \cap R|$ and $|\Phi \cap S|$ are independent.

On many occasions, it is convenient to consider Φ conditioned to have a point located at $z \in \mathcal{D}$ and we let Φ_z be the corresponding random point set. Classical results on Poisson point processes ensure that $\Phi_z \setminus \{z\}$ is distributed like Φ , so that one can take $\Phi_z = \Phi \cup \{z\}$, for Φ independent of z [see, e.g., 1, Section 1.4].

3.1 Preliminaries

We establish the following notation (see Figure 2). Let $\mathbf{Z} = (Z_i, i > 0)$ denote the sequence of stoppers visited during the walk with $Z_0 := z$. Let $L_i = \|Z_i q\|$ denote the distance to the destination q . The distance L_i is strictly decreasing and the point set Φ is almost surely finite, thus ensuring that the walk stops after a finite number of steps κ , at which point we have $Z_\kappa = q$. For $x > 0$, we also let $\kappa(x)$ be the number of steps required to reach a point within distance x of the query. Thus, using the fact that L_i is decreasing, we get that $i < \kappa(x)$ if and only if $L_i > x$. The important parameters needed to track the location and progress of the walk are the radius R_i such that $Z_{i+1} \in \partial\text{Cone}(Z_i, q, R_i)$, and the angle α_i between $Z_i q$ and $Z_i Z_{i+1}$. $\text{Disc}(Z_i, q, R_i)$ may contain several points of Φ , let τ_i denote $|\text{Disc}(Z_i, q, R_i) \setminus \{Z_i, Z_{i+1}\} \cap \Phi|$ the number of such points and N_i the number of these points along with their Delaunay neighbours.

In order to compute the walk efficiently, the algorithm presented gathers a lot of information. In particular, we access all the points in $\text{Disc}(Z_i, q, R_i)$ and their neighbours. For the analysis, we want to keep the landscape as concise as possible, and so we define a filtration which only

contains the necessary information for the walk to be a measurable process. Let \mathcal{F}_i denote the information consisting of (the σ -algebra generated by) the locations of the points of Φ contained in $\cup_{j=0}^i \text{Disc}(Z_j, q, R_j)$. Finally, we shall write ω_n to denote a sequence satisfying $\omega_n \geq \log n$. (We do not really need an upper bound on ω_n , but of course, if it is too large, some of the results become vacuus; one could typically think of it as $\log n$ or a polylog.)

We often need to condition on the size of the largest empty ball within the process Φ_n . This is dealt with in the following lemma.

Lemma 10. *Let $b(\mathbf{x}, r)$ denote the closed ball of radius r centred at \mathbf{x} . Then $\forall c > 0, \xi > 0$,*

$$\mathbb{P}(\exists x \in \mathcal{D}_n : b(x, c\omega_n^{1/2+\xi}) \cap \Phi_n = \emptyset) \leq \exp(-\omega_n^{1+\xi})$$

for n sufficiently large.

Proof. We have

$$\mathbb{P}(\exists x \in \mathcal{D}_n : b(x, c\omega_n^{1/2+\xi}) \cap \Phi_n = \emptyset) \leq \mathbb{P}(\exists B \in P : B \cap \Phi_n = \emptyset),$$

where P is any maximal packing of \mathcal{D}_n with balls B of radius $\frac{1}{2}c\omega_n^{1/2+\xi}$ centred in \mathcal{D}_n . (Here, by packing, we mean that one could not add any ball to the collection, but we do not enforce the balls to be fully contained in \mathcal{D}_n , only that their centers lie in \mathcal{D}_n ; so in particular, such a packing is never empty, even if the balls are larger than \mathcal{D}_n .) If the radius of curvature of \mathcal{D}_n is lower bounded by $c\omega_n^{1/2+\xi}$ (which happen for n large enough) such a ball B contains a ball of radius $\frac{1}{4}c\omega_n^{1/2+\xi}$ entirely inside \mathcal{D}_n . For n large enough, any such packing contains at most n balls and we have

$$\mathbb{P}(\exists x \in \mathcal{D}_n : b(x, c\omega_n^{1/2+\xi}) \cap \Phi_n = \emptyset) \leq n \exp(-\pi \frac{1}{4^2}) c^2 \omega_n^{1+2\xi} \leq \exp(-\omega_n^{1+\xi}). \quad \square$$

3.1.1 The size of the discs

If the search cone $\text{Cone}(Z_i, q, \infty)$ does not intersect any of the previous discs, the region which determines R_{i+1} is ‘fresh’ and R_{i+1} is independent of \mathcal{F}_i . Lemma 7 provides a condition which guarantees independence of the search cones. To take advantage of it, we write $\xi := 2 + \sqrt{2}$, and for $i \geq 0$, define the event

$$G_i := \{\forall j \leq i + 1, R_j < \omega_n/\xi\}, \quad (1)$$

Then if the event $G_i^* := G_i \cap \{L_i \geq \omega_n\}$ occurs; for every $j \leq i$, the search-cone $\text{Cone}(Z_j, q, \infty)$ does not intersect any of the regions $\text{Disc}(Z_k, q, R_k)$, $0 \leq k < j$, and the corresponding variables (R_j, α_j) , $0 \leq j \leq i + 1$ are independent. Although it might seem like an odd idea, G_i^* does include some condition on R_{i+1} ; this ensures that on G_i^* , we have $L_{i+1} > L_i - 2R_{i+1} > 0$, so that $i + 1$ is not the last step. So for $x > 0$ we have

$$\begin{aligned} \mathbb{P}(R_{i+1} > x \mid \mathcal{F}_i, G_i^*) &= \mathbb{P}(\Phi \cap \text{Cone}(Z_i, q, x) \setminus \{Z_i\} = \emptyset \mid \mathcal{F}_i, G_i^*) \mathbf{1}_{\{\xi x \leq \omega_n\}} \\ &= \exp(-Ax^2) \mathbf{1}_{\{\xi x \leq \omega_n\}}, \end{aligned} \quad (2)$$

where A denotes the area of $\text{Cone}(z, q, 1)$ which is the shaded region in Figure 1. Indeed, conditional on \mathcal{F}_i and G_i^* , $|\Phi \cap \text{Cone}(Z_i, q, x) \setminus \{Z_i\}|$ is a Poisson random variable with mean Ax^2 where

$$A := 2 \left(\cos \frac{\pi}{8} \sin \frac{\pi}{8} + \frac{\pi}{8} \right) = \frac{\sqrt{2}}{2} + \frac{\pi}{4}. \quad (3)$$

We will repeatedly use the conditioning on G_i to introduce independence, and it is important to verify that G_i indeed occurs with high probability. For G_i to fail, there must be a first step j for which $R_j \geq \omega_n/\xi$. Writing G_i^c for the complement of G_i and defining G_{-1} to be a void conditioning: provided that $i = O(n)$ (which will always be the case in the following)

$$\begin{aligned} \mathbb{P}(G_i^c) &\leq \sum_{0 \leq j \leq i+1} \mathbb{P}(R_j \geq \omega_n/\xi \mid G_{j-1}) \\ &\leq \exp(\log O(n) - A\omega_n^2/\xi^2) \\ &\leq \exp(-\omega_n^{3/2}) \end{aligned} \quad (4)$$

for all n large enough since $\omega_n \geq \log n$.

Remark about the notation. It is convenient to work with an “ideal” random variable that is not constrained by the location of the query or artificially forced to be at most ω_n/ξ , and we define \mathcal{R} by $\mathbb{P}(\mathcal{R} \geq x) = \exp(-Ax^2)$ for $x \geq 0$. In the course of the proof, we use multiple other such ideal random variables, to distinguish them from the ones arising from the actual process, we use calligraphic letters to denote them.

3.1.2 The progress for one step

We now focus on the distribution of the angle $\angle qZ_iZ_{i+1}$ and by extension the progress made during one step in the walk. Let $\text{Cone}_\alpha(z, q, r)$ be the cone of half angle α with the same apex and axis as $\text{Cone}(z, q, r)$. For $S \subset \mathbb{R}^2$, let $\mathcal{A}(S)$ denote its area. On the event G_i^* , $Z_{i+1} \neq q$ and α_{i+1} is truly random and its distribution is symmetric and given by (see Figure 7):

$$\begin{aligned} \mathbb{P}(|\alpha_{i+1}| < x \mid R_{i+1} = r, \mathcal{F}_i, G_i^*) &= \lim_{\varepsilon \rightarrow 0} \frac{\mathcal{A}(\text{Cone}_x(Z_i, q, r + \varepsilon) \setminus \text{Cone}_x(Z_i, q, r))}{\mathcal{A}(\text{Cone}(Z_i, q, r + \varepsilon) \setminus \text{Cone}(Z_i, q, r))} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{((r + \varepsilon)^2 - r^2)(x + \frac{1}{2} \sin 2x)}{((r + \varepsilon)^2 - r^2)(\frac{\pi}{8} + \frac{\sqrt{2}}{4})} \\ &= \frac{8}{\pi + 2\sqrt{2}} \left(x + \frac{\sin 2x}{2} \right). \end{aligned} \quad (5)$$

So in particular, conditional on \mathcal{F}_i and G_i^* , α_{i+1} is independent of R_{i+1} . We will write α for the ‘ideal’ angle distribution given by (5), and enforce that \mathcal{R} and α be independent.

3.2 Geometric and combinatorial parameters

In this section we will build the elements required to bound the algorithmic complexity of the CONE-WALK algorithm. We begin by bounding the number of *steps* (or equivalently, the number of *stoppers*) required by the walk process to reach the destination. We will then bound the

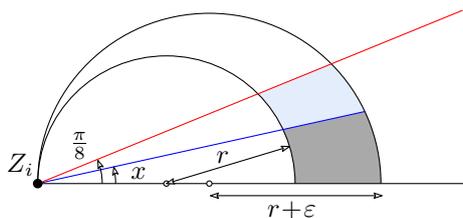


Figure 7: For the angle to be smaller than x given $R_{i+1} \in [r, r + \varepsilon]$, the stopper must fall within the dark shaded region

number of vertices *visited* by the walk process, recalling that this will involve bounding the number of *intermediary vertices* within the discs $\text{Disc}(Z_i, q, R_i)$ at each step. The final part of the proof will be to bound the number of vertices *accessed* by the CONE-WALK algorithm when constructing the sequence of stoppers and intermediary vertices. The vertices *accessed* will include all of the vertices visited, and their 1-hop neighbourhood.

3.2.1 The maximum number of vertices accessed during a step

At each step during a walk, we do not a priori access a bounded number of sites when performing a search for the next stopper. Such a bound is important to limit the number of neighbourhoods that may be accessed during one step, since we note that the maximum number of vertices accessed during one step explicitly provides an upper bound on the number of neighbourhoods accessed. An easy bound of $\log^{1+\varepsilon} n$, for any $\varepsilon > 0$ may be obtained when considering pairs of start and destination points at least $\sqrt{\log n}$ away from the border of $\partial\mathcal{D}$. However, we opt to explicitly take care of border effects, giving us a slightly weaker bound that can be applied everywhere.

Proposition 11. *Let M_{\max} be the maximum number of vertices accessed during any step in any walk. Then*

$$\mathbb{P}\left(M_{\max} \geq \omega_n^{3+\xi}\right) \leq 2 \exp\left(-\omega_n^{1+\xi/4}\right).$$

In the following, we note that M_{\max} is bounded by $\tau_{\max} \cdot \Delta_{\Phi}$, where Δ_{Φ} gives the maximum degree of any vertex contained within $\text{DT}(\Phi)$ and τ_{\max} is the maximum number sites contained within any step in any instance of cone walk. We thus focus on bounding τ_{\max} , and our result will follow directly from the proof of Proposition 30 in Appendix A.

Lemma 12.

$$\mathbb{P}(\tau_{\max} > \omega_n^{1+\xi}) \leq \exp\left(-\omega_n^{1+\xi/3}\right).$$

Proof. Let A be the event that the maximum disc radius for any step in any walk is bounded by $\frac{1}{2}\omega_n^{1/2+\xi}$ and let B be the event that every ball $b(x, \frac{1}{2}\omega_n^{1/2+\xi})$ contains fewer than $\omega_n^{1+2\xi}$ points of Φ , for $x \in \mathcal{D}$. We have, for n large enough,

$$\begin{aligned} \mathbb{P}(\tau_{\max} > \omega_n^{1+2\xi}) &\leq \mathbb{P}(\tau_{\max} > \omega_n^{1+2\xi} \mid A \cap B) + \mathbb{P}(A^c) + \mathbb{P}(B^c) \\ &\leq \exp\left(-\omega_n^{1+\xi}\right) + \exp\left(-\omega_n^{1+\xi}\right). \end{aligned}$$

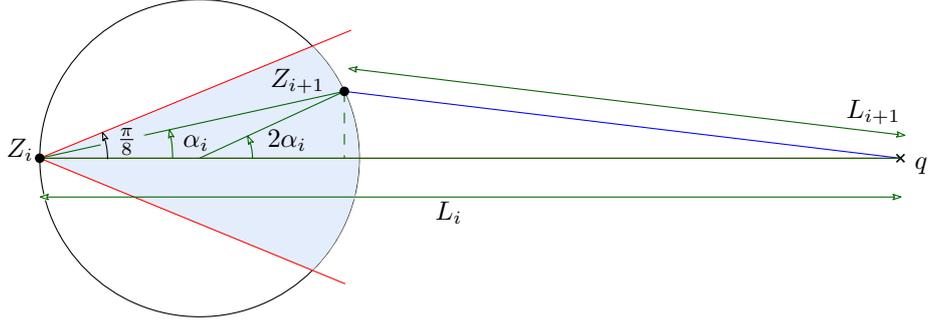


Figure 8: Computing distance progress at step i .

Note that the bound on $\mathbb{P}(A^c)$ is implied by Lemma 10 since a large disc implicitly has a large empty cone. For the bound on $\mathbb{P}(B^c)$, we imagine splitting \mathcal{D} into a uniform grid with squares of side $\frac{1}{2}\omega_n^{1+\xi}$. The proof follows by noting that every ball of radius $\frac{1}{2}\omega_n^{1+\xi}$ is contained in a group of at most four adjacent squares, each of which must contain at least $\frac{1}{4}\omega_n^{1+2\xi}$ sites. We then use the fact that $n\mathbb{P}(\text{Po}(\frac{1}{4}\omega_n^{1+\xi}) \geq \frac{1}{4}\omega_n^{1+2\xi}) \leq \exp(-\omega_n^{1+\xi})$ for n large enough. We omit the details. \square

3.2.2 The number of steps in the walk

Recall that a new step is defined each time a new stopper is visited. We will start with a first crude estimate for the decrease in distance after a given number of steps. Note that $L_i = \|Z_i q\|$ and α_i denotes the angle between $Z_i Z_{i+1}$ and $Z_i q$. Simple geometry implies (see Figure 8):

$$\begin{aligned} L_i - R_i(1 + \cos(2\alpha_i)) &\leq L_{i+1} = \sqrt{(L_i - R_i(1 + \cos(2\alpha_i)))^2 + R_i^2 \sin^2(2\alpha_i)} \\ &\leq L_i - R_i(1 + \cos(2\alpha_i)) + 2\frac{R_i^2}{L_i}, \end{aligned} \quad (6)$$

since $\sqrt{1-x} \leq 1 - x/2$ for any $x \in [0, 1]$. As a consequence

$$L_0 - \sum_{j=0}^{i-1} R_j(1 + \cos(2\alpha_j)) \leq L_i \leq L_0 - \sum_{j=0}^{i-1} R_j(1 + \cos(2\alpha_j)) + \frac{2}{\omega_n} \cdot \sum_{j=0}^{i-1} R_j^2. \quad (7)$$

In particular, since $\omega_n \rightarrow \infty$, after i steps, the expected distance $\mathbb{E}[L_i]$ to the aim q should not be far from $L_0 - i\mathbb{E}[R(1 + \cos(2\alpha))]$. Furthermore, conditional on G_i , and for i such that $L_i \geq \omega_n$, the summands involved in Equation (7) are independent, bounded by $2\omega_n$ and have bounded variance, so that the sum should be highly concentrated about its expected value [9, 15, 21]. In other words, one expects that for i much larger than $L_0/\mathbb{E}[R(1 + \cos 2\alpha)]$, it should be the case that $L_i \leq \omega_n$ with fairly high probability. Making this formal constitutes the backbone of our proof.

Lemma 13. *Let $z \in \mathcal{D}$, suppose that $\ell \geq 1$ is such that $L_0 = \|zq\| \geq (\ell + 1)\omega_n$. Consider $\text{DT}(\Phi_z)$. There exists a constant $\eta > 0$ such that*

$$\mathbb{P}(L_0 - L_\ell \leq \ell \mathbb{E}[\mathcal{R}]/2) \leq \exp(-\eta\ell/\omega_n) + \exp(-\omega_n^{3/2}).$$

Proof. We use the crude bounds $R_i \leq L_i - L_{i+1} \leq 2R_i$ (see Figure 8). It follows that

$$\begin{aligned} \mathbb{P}(L_0 - L_\ell \leq \ell \mathbb{E}[\mathcal{R}]/2) &\leq \mathbb{P}(L_0 - L_\ell \leq \ell \mathbb{E}[\mathcal{R}]/2 \mid G_\ell) + \mathbb{P}(G_\ell^c) \\ &\leq \mathbb{P}\left(\sum_{j=0}^{\ell-1} R_j \leq \frac{\ell}{2} \mathbb{E}[\mathcal{R}] \mid G_\ell\right) + \exp(-\omega_n^{3/2}), \end{aligned}$$

by (4), since the constraint on ℓ imposes that $\ell = O(\sqrt{n})$. Now, since $L_0 \geq (\ell + 1)\omega_n$ and $\xi > 2$, on the event G_ℓ , we have $L_i \geq \omega_n$ for $0 \leq i \leq \ell$ so that G_ℓ^* occurs: conditional on G_ℓ , the search cones do not intersect and the random variables R_j , $0 \leq j \leq \ell$ are independent and identically distributed (see Lemma 7). Furthermore, we have

$$\begin{aligned} \mathbb{E}[R_j \mid G_\ell] &= \int_0^\infty \mathbb{P}(R_j \geq x \mid G_\ell) dx \\ &\geq \int_0^{\omega_n/\xi} \exp(-Ax^2) dx \\ &\geq \mathbb{E}[\mathcal{R}] - \exp(-\omega_n^{3/2}), \end{aligned}$$

for all n large enough. It follows that for all n large enough, by Theorem 2.7 of [21, p. 203]

$$\begin{aligned} \mathbb{P}\left(\sum_{j=0}^{\ell-1} R_j \leq \frac{\ell}{2} \mathbb{E}[\mathcal{R}] \mid G_\ell\right) &\leq \mathbb{P}\left(\sum_{j=0}^{\ell-1} (R_j - \mathbb{E}[R_j \mid G_\ell]) \leq -\frac{\ell}{3} \mathbb{E}[R_0 \mid G_\ell] \mid G_\ell\right) \\ &\leq \exp\left(-\frac{t^2}{2\ell \mathbb{V}(R_0 \mid G_\ell) + 2t\omega_n/3}\right) \quad t = \ell \mathbb{E}[R_0 \mid G_\ell]/3 \\ &\leq \exp(-\eta\ell/\omega_n), \end{aligned}$$

for some constant $\eta > 0$ independent of ℓ and n . \square

The rough estimate in Lemma 13 may be significantly strengthened, and the very representation in (7) yields a bound on the number of search cones or steps that are required to get within distance ω_n of the query point q . (If the starting site z satisfies $L_0 = \|zq\| \leq \omega_n$, then this phase does not contain any step.)

Proposition 14. *Let $z \in \mathcal{D}_n$, and let $\kappa(\omega_n)$ denote the number of steps of the walk to reach a site which is within distance ω_n of q in $\Phi_z \cup \{q\}$ when starting from the site $z \in \Phi_z$ at distance $L_0 = \|zq\| \geq \omega_n$. Then*

$$\mathbb{P}\left(\left|\kappa(\omega_n) - \frac{L_0}{\mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)]}\right| \geq 2\omega_n^2 \sqrt{2L_0} + \omega_n\right) \leq 4 \exp(-\omega_n^{3/2}).$$

Proof. We now make formal the intuition that follow Equation (7). We start with the upper bound. For any integer $k \geq 0$, we have

$$\begin{aligned}\mathbb{P}(\kappa(\omega_n) \geq k) &= \mathbb{P}(L_k \geq \omega_n) \\ &\leq \mathbb{P}(L_k \geq \omega_n | G_k) + \mathbb{P}(G_k^c),\end{aligned}$$

and since the second term is bounded in (4), it now suffices to bound the first one. However, given G_k and $L_k \geq \omega_n$, the random variables (R_i, α_i) , $i = 1, \dots, k$ are independent and identically distributed. The only effect of this conditioning is that R_i is distributed as \mathcal{R} conditioned on $\mathcal{R} < \omega_n/\xi$.

Write $X_i = R_i(1 + \cos 2\alpha_i) - 2R_i^2/\omega_n$, and note that $X_i \geq 0$ if $R_i \leq \omega/\xi$. Then, from (7), we have

$$\mathbb{P}(L_k \geq \omega_n | G_k) \leq \mathbb{P}\left(\sum_{i=0}^{k-1} X_i \leq L_0 - \omega_n \mid G_k, L_k \geq \omega_n\right).$$

Conditional on $G_k^* = G_k \cap \{L_k \geq \omega_n\}$, the random variables X_i are independent, $0 \leq X_i \leq 2R_i \leq \omega_n$. Furthermore, since X_i has Gaussian tails, its variance (conditional on G_k) is bounded by a constant independent of i and n . Choosing $k_0 = \lceil (L_0 + t)/\mathbb{E}[X_0 | G_0^*] \rceil$, for some $t < L_0$ to be chosen later, and using the Bernstein-type inequality in Theorem 2.7 of [21, p. 203], we obtain

$$\begin{aligned}\mathbb{P}(L_{k_0} \geq \omega_n | G_{k_0}) &\leq \mathbb{P}\left(\sum_{i=0}^{k_0-1} (X_i - \mathbb{E}[X_i | G_{k_0}]) \leq -t \mid G_{k_0}^*\right) \\ &\leq \exp\left(-\frac{t^2}{2k_0\mathbb{V}(X_0 | G_0^*) + 2\omega_n t/3}\right).\end{aligned}$$

In particular, for $t = \omega_n^3\sqrt{L_0}$, we have for all n large enough $\mathbb{P}(L_{k_0} \geq \omega_n | G_{k_0}) \leq \exp(-\omega_n^2)$, since $L_0 \geq \omega_n$.

A matching lower bound on $\kappa(\omega_n)$ may be obtained similarly, using the lower bound on L_{i+1} in Equation (6) and following the approach we used to devise the upper bound with $X'_i = R_i(1 + \cos 2\alpha_i)$ (we omit the details). It follows that, for $k_1 = \lfloor (L_0 + t)/\mathbb{E}[X'_0 | G_0^*] \rfloor$, we have $\mathbb{P}(L_{k_1} \leq \omega_n | G_{k_1}) \leq \exp(-\omega_n^2)$.

To complete the proof, it suffices to estimate the difference between k_0 and k_1 . We have

$$\begin{aligned}\mathbb{E}[X_0 | G_0^*] &= \mathbb{E}[R_0(1 + \cos 2\alpha_0) | G_0^*] - \frac{2\mathbb{E}[R_0^2 | G_0^*]}{\omega_n} \\ &= \mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)] + O(1/\omega_n),\end{aligned}$$

and similarly, $\mathbb{E}[X'_0 | G_0^*] = \mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)]$. It follows that $|k_1 - k_0| = O(L_0/\omega_n)$, which is not strong enough to prove the claim. So we need to strengthen the upper bound on the second sum in the right-hand side of (7). We quickly sketch how to obtain the required estimate. The idea is to use a dyadic argument to decompose $\kappa(\omega_n)$ into the number of steps to reach $L_0/2^j$,

for $j \geq 1$, until one gets to ω_n for $j = j_0 := \lceil \log_2(L_0/\omega_n) \rceil$. For the steps i which are taken from Z_i with $L_i/L_0 \in (2^{-j}, 2^{-j+1}]$, we use the improved bound

$$L_{i+1} \leq L_i - R_i(1 + \cos 2\alpha) + \frac{2}{L_0 2^{-j}} R_i^2.$$

Then write

$$\kappa(\omega_n) = \sum_{j=1}^{j_0} [\kappa(L_0/2^j) - \kappa(L_0/2^{j-1})],$$

and observe that the j -th summand is stochastically dominated by $\kappa(L'_0/2)$ where $L'_0 = L_0/2^{j-1}$. For each j , we define $k_0(j) = \lceil (L_0/2^j + t_j)/\mathbb{E}[X_0 | G_0^*] \rceil$ where $t_j := \omega_n^2 \sqrt{L_0/2^j}$ and note that

$$\begin{aligned} \sum_{j=1}^{j_0} k_0(j) &\leq \frac{1}{\mathbb{E}[X_0 | G_0^*]} \sum_{j=1}^{j_0} (L_0/2^j + t_j) + \lceil \log_2(L_0/\omega_n) \rceil \\ &\leq \frac{L_0}{\mathbb{E}[X_0 | G_0^*]} + 2\omega_n^2 \sqrt{2L_0} + \omega_n, \end{aligned}$$

for $\omega_n \geq \log n$, since $\pi L_0^2 \leq n$. In other words, if $\kappa(L_0/2^j) - \kappa(L_0/2^{j-1}) \leq k_0(j)$ for every j , then $\kappa(\omega) \leq L_0/\mathbb{E}[X_0 | G_0^*] + 2\omega_n^2 \sqrt{2L_0} + \omega_n$. The claim follows easily by using the union bound, where in each stretch $[L_0/2^j, L_0/2^{j-1})$ we bound the number of steps using the previous arguments. \square

Corollary 15. *Let $z \in \mathcal{D}_n$, and let κ denote the number of steps of the walk to reach the objective q in $\Phi_z \cup \{q\}$ when starting from the site $z \in \Phi_z$ at distance $L_0 = \|zq\|$. Then*

$$\mathbb{P} \left(\kappa > \frac{L_0}{\mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)]} + 2\omega_n^2 \sqrt{2L_0} + \omega_n^3 \right) \leq 5 \exp \left(-\omega_n^{3/2} \right).$$

Proof. It suffices to bound the number of steps i such that $L_i < \omega_n$. Since L_i is decreasing, the walk only stops at most once at any given site, and the number of steps i with $L_i \leq \omega_n$ is at most the number of sites lying within distance ω_n of q . Recalling that $\text{Po}(x)$ denotes a Poisson random variable with mean x . We have [15]

$$\begin{aligned} \mathbb{P} (\#\{i < \kappa : L_i \leq \omega_n\} \geq 2\pi\omega_n^2) &\leq \mathbb{P} (\text{Po}(\pi\omega_n^2) \geq 2\pi\omega_n^2) \\ &\leq \exp(-\pi\omega_n^2/3). \end{aligned}$$

The claim then follows easily from the upper bound in Proposition 14. \square

3.2.3 The number of vertices in the discs

We now bound the total number of vertices visited, which we recall is exactly the number of points in Φ_n falling within the union of all of the discs in the walk. Proposition 14 will be the key to analysing the path constructed by the walk: representations based on sums of random variables similar to the one in (6) may be obtained to upper bound the number of steps and intermediate steps visited by the walk (which is an upper bound on the vertices visited by the path), and also the sum of the length of the edges.

Proposition 16. *Let $K = K(z)$ be the number of vertices visited by the walk starting from a given site z with $L_0 = \|zq\|$. Then, for all n large enough,*

$$\mathbb{P}\left(K \geq \frac{L_0}{\mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)]} \cdot \frac{\pi - A}{A} + \sqrt{L_0\omega_n^4 + \omega_n^3}\right) \leq 7 \exp\left(-\omega_n^{3/2}\right).$$

Proof. There are two contributions to $K - \kappa(\omega_n)$: first the number of intermediate steps which lie at distance greater than ω_n from q , and all the sites which are visited and lie within distance ω_n from q . Let $K = K_1 + K_2$ where K_1 and K_2 denote these two contributions, respectively. By the proof of Corollary 15, we have

$$\mathbb{P}(K_2 \geq 2\pi\omega_n^2) \leq \exp(-\pi\omega_n^2/3). \quad (8)$$

To bound K_1 , observe that the monotonicity of L_i implies that K_1 counts precisely the number of intermediate steps before reaching the disc of radius ω_n about q . Observe that if $L_0 < \omega_n$, $K_1 = 0$, so we may assume that $L_0 \geq \omega_n$. Recall that τ_i denotes the number of *intermediate* points at the i -th step. Note that the intermediate points counted by τ_i all lie in $\text{Disc}(Z_i, q, R_i) \setminus \text{Cone}(Z_i, q, R_i)$, and given the radius R_i , τ_i is stochastically bounded by a Poisson random variable with mean $(\pi - A)R_i^2$. Furthermore, on the event $G_{\kappa(\omega_n)}$, the random variables R_i , $i = 0, \dots, \kappa(\omega_n)$ are independent. Also, by Lemma 8 the regions $\text{Disc}(Z_i, q, R_i) \setminus \text{Cone}(Z_i, q, R_i)$, $i \geq 0$, are disjoint so that the random variables τ_i , $i = 0, \dots, \kappa$ are independent given R_i , $i = 0, \dots, \kappa$.

Let $\tilde{\mathcal{R}}_i$, $i \geq 0$, be a sequence of i.i.d. random variables distributed like \mathcal{R} conditioned on $\mathcal{R} \leq \omega_n/\xi$ and given this sequence, let $\tilde{\tau}_i$, $i \geq 0$, be independent distributed like $\text{Po}((\pi - A)\tilde{\mathcal{R}}_i)$. As a consequence of the previous arguments, for $k = k_0 + 2\omega_n^2\sqrt{2L_0} + \omega_n$ with $k_0 = \lceil L_0/\mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)] \rceil$, we have

$$\begin{aligned} \mathbb{P}(K_1 \geq \ell) &\leq \mathbb{P}\left(\sum_{i=0}^{(k-1) \wedge \kappa(\omega_n)} \tau_i \geq \ell\right) + \mathbb{P}\left(\kappa(\omega_n) \geq k_0 + 2\omega_n^2\sqrt{2L_0} + \omega_n\right) \\ &\leq \mathbb{P}\left(\sum_{i=0}^{k-1} \tilde{\tau}_i \geq \ell\right) + \mathbb{P}(G_k^c) + \mathbb{P}\left(\kappa(\omega_n) \geq k_0 + 2\omega_n^2\sqrt{2L_0} + \omega_n\right) \\ &\leq \mathbb{P}\left(\sum_{i=0}^{k-1} \tilde{\tau}_i \geq \ell\right) + 5 \exp\left(-\omega_n^{3/2}\right), \end{aligned} \quad (9)$$

by (4) and Proposition 14.

We now bound the first term in (9). Note that $i \geq 0$, we have

$$\mathbb{E}[\tilde{\tau}_i] = (\pi - A)\mathbb{E}[\tilde{\mathcal{R}}_i^2] \leq (\pi - A)\mathbb{E}[\mathcal{R}^2] = \frac{\pi - A}{A} =: \gamma,$$

and we expect that $\sum_{i=0}^{k-1} \tilde{\tau}_i$ should not exceed its expected value, $k\gamma$ by much. Write $\ell = k\gamma + t$, for some t to be chosen later. For the sum to be exceptionally large either the radii of the search

discs are large, or the discs are not too large but the number of points are:

$$\begin{aligned} \mathbb{P}\left(\sum_{i=0}^{k-1} \tilde{\tau}_i \geq k\gamma + t\right) &= \mathbb{P}\left(\text{Po}\left((\pi - A) \sum_{i=0}^{k-1} \tilde{\mathcal{R}}_i^2\right) \geq k\gamma + t\right) \\ &\leq \mathbb{P}\left(\text{Po}\left(k\gamma + \frac{t}{2}\right) \geq k\gamma + t\right) + \mathbb{P}\left(\sum_{i=0}^{k-1} \tilde{\mathcal{R}}_i^2 \geq \frac{k\gamma + t/2}{\pi - A}\right). \end{aligned} \quad (10)$$

The first term simply involves tail bounds for Poisson random variables. For $t = \sqrt{L_0}\omega_n^4$, we have

$$\begin{aligned} \mathbb{P}\left(\text{Po}\left(k\gamma + \frac{t}{2}\right) \geq k\gamma + t\right) &\leq \exp\left(-\frac{(t/2)^2}{3(k\gamma + t/2)}\right) \\ &\leq \exp(-\omega_n^3), \end{aligned}$$

for n large (recall that we can assume here that $L_0 \geq \omega_n$.) The second term in (10) is bounded using the same technique as in the proof of Proposition 14 above. Since we have $0 \leq \tilde{\mathcal{R}}_i^2 \leq \omega_n^2$ and $\mathbb{E}[\tilde{\mathcal{R}}_i^2] \leq 1/A$, we obtain for some positive constant c ,

$$\mathbb{P}\left(\sum_{i=0}^{k-1} \tilde{\mathcal{R}}_i^2 \geq \frac{k\gamma + t/2}{\pi - A}\right) \leq \mathbb{P}(G_k) + \exp\left(-8\frac{k}{A^2\omega_n^4}\right) \leq \exp\left(-\frac{ct^2}{k\mathbb{V}(\mathcal{R}) + \omega_n^2 t}\right).$$

Recalling that $L_0 \geq \omega_n$, yields

$$\mathbb{P}\left(\sum_{i=0}^{k-1} \tilde{\tau}_i \geq k\gamma + \omega_n^4 \sqrt{L_0}\right) \leq \exp(-\omega_n^2),$$

for all n large enough, which together with (9) proves that $\mathbb{P}(K_1 \geq k\gamma + \omega_n^4 \sqrt{L_0}) \leq 6 \exp(-\omega_n^{3/2})$. Using (8) readily yields the claim. \square

3.2.4 The length of SIMPLE-PATH

When using COMPETITIVE-PATH, the path length is deterministically bounded by the length of the walk. However, for SIMPLE-PATH, the path length is dependent on the configuration of the points inside the discs. We will show that with strong probability and as long as the walk is sufficiently long, the path length given by COMPETITIVE-PATH is no better in an asymptotic sense than that given by SIMPLE-PATH.

Proposition 17. *For $z \in \mathcal{D}$, let $\Lambda = \Lambda(z)$ be the sum of the lengths of the edges of $\text{DT}(\Phi_z)$ used by SIMPLE-PATH given a walk with objective q and starting from z such that $L_0 = \|zq\|$. Then,*

$$\mathbb{P}\left(\Lambda \geq cL_0 + (3\sqrt{L_0} + 1)\omega_n^4\right) \leq 8 \exp(-\omega_n^{3/2}) \quad \text{where} \quad c := \frac{22\pi - 4\sqrt{2}}{2 + 3\pi + 8\sqrt{2}}. \quad (11)$$

Proof. Write λ_i for the sum of the lengths of the edges used by the walk to go from Z_i to Z_{i+1} . So $\Lambda = \sum_{i=0}^{\kappa-1} \lambda_i$. Our bound here is very crude: all the intermediate points remain in $\text{Disc}(Z_i, q, R_i)$, and given R_i , we have $\lambda_i \leq (1 + \tau_i) \cdot 2R_i$. Again, on G_k the cones do not intersect provided that $L_k \geq \omega_n$, and by Lemma 8 the random variables λ_i , $0 \leq i < k$ are independent. We use once again the method of bounded variances (Theorem 2.7 of [21]).

We decompose the sum into the contribution of the steps before $\kappa(\omega_n)$ and the ones after:

$$\mathbb{P}(\Lambda \geq x + t) \leq \mathbb{P}\left(\sum_{i=0}^{(k-1) \wedge \kappa(\omega_n)} \lambda_i \geq x\right) + \mathbb{P}(\kappa(\omega_n) \geq k) + \mathbb{P}\left(\sum_{i=\kappa(\omega_n)}^{\kappa-1} \lambda_i \geq t\right). \quad (12)$$

For $i \geq \kappa(\omega_n)$, $\text{Disc}(Z_i, q, R_i)$ is contained in $b(q, \omega_n)$, the disc of radius ω_n around q , and the contribution of the steps $i \geq \kappa(\omega_n)$ is at most $2\omega_n |\Phi \cap b(q, \omega_n)|$. In particular

$$\begin{aligned} \mathbb{P}\left(\sum_{i=\kappa(\omega_n)}^{\kappa-1} \lambda_i \geq t\right) &\leq \mathbb{P}(2\omega_n \text{Po}(\pi\omega_n^2) \geq t) \\ &\leq \exp\left(-\omega_n^{3/2}\right), \end{aligned}$$

for all n large enough provided that $t \geq 4\pi\omega_n^3$. To make sure that the second contribution in (12) is also small, we rely on Proposition 14 and choose $k = \lceil L_0/\mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)] + 2\omega_n^2\sqrt{2L_0} + \omega_n \rceil$ so that $\mathbb{P}(\kappa(\omega_n) \geq k) \leq 4 \exp\left(-\omega_n^{3/2}\right)$.

Finally, to deal with the first term in (12), we note that on G_k , the random variables λ_i , $i = 0, \dots, k+1$ are independent given R_i , $i = 0, \dots, k+1$. Let $\tilde{\mathcal{R}}_i$, $i = 0, \dots, k-1$ be i.i.d. copies of \mathcal{R} conditioned on $\mathcal{R} \leq \omega_n/\xi$; then let $\tilde{\tau}_i$ be independent given R_i , $i = 0, \dots, k+1$, and such that $\tilde{\tau}_i = \text{Po}((\pi - A)\tilde{\mathcal{R}}_i)$; finally, let $\tilde{\lambda}_i = 2R_i(1 + \tilde{\tau}_i)$. We choose $x = k\mathbb{E}[\tilde{\lambda}_0] + y$ with $y = \sqrt{L_0}\omega_n^4$. Using arguments similar to the ones we have used in the proofs of Propositions 14 and 16, we obtain

$$\begin{aligned} \mathbb{P}\left(\sum_{i=0}^{(k-1) \wedge \kappa(\omega_n)} \lambda_i \geq x\right) &\leq \mathbb{P}\left(\sum_{i=0}^{k-1} \tilde{\lambda}_i \geq x\right) + \mathbb{P}(G_k^c) \\ &\leq \exp\left(-\frac{y^2}{2k\mathbb{V}(\tilde{\lambda}_0) + 2\omega_n^2 y/3}\right) + \mathbb{P}\left(\exists i < k : \tilde{\lambda}_i \geq \omega_n^2\right) + \mathbb{P}(G_k^c). \end{aligned} \quad (13)$$

To bound the second term in the right-hand side above, observe that for $i < k$, we have, for any $x > 0$,

$$\begin{aligned} \mathbb{P}(\tilde{\lambda}_i \geq x^2) &\leq \mathbb{P}((1 + \tilde{\tau}_i)2\tilde{\mathcal{R}}_i \geq x^2) \\ &\leq \mathbb{P}((1 + \tilde{\tau}_i)2\mathcal{R}_i \geq x^2 \mid 2\tilde{\mathcal{R}}_i \leq x) + \mathbb{P}(2\tilde{\mathcal{R}}_i \geq x) \\ &\leq \mathbb{P}(1 + \tilde{\tau}_i \geq x \mid 2\tilde{\mathcal{R}}_i \leq x) + \mathbb{P}(2\tilde{\mathcal{R}}_i \geq x) \\ &\leq \mathbb{P}(\text{Po}((\pi - A)x^2/4) \geq x - 1) + \exp(-Ax^2/4) \\ &\leq 2 \exp(-\eta x^2), \end{aligned}$$

for some constant $\eta > 0$ and all x large enough. It follows immediately that $\mathbb{V}(\tilde{\lambda}_0) < \infty$ and that, n large enough,

$$\begin{aligned} \mathbb{P}(\exists i < k : \tilde{\lambda}_i \geq \omega_n^2) &\leq k \exp(-\eta\omega_n^2) \\ &\leq \exp(-\omega_n^{3/2}). \end{aligned} \tag{14}$$

Going back to (13), we obtain

$$\mathbb{P}\left(\sum_{i=0}^{(k-1) \wedge \kappa(\omega_n)} \lambda_i \geq x\right) \leq 3 \exp(-\omega_n^{3/2}),$$

since here, we can assume that $L_0 \geq \omega_n$ (if this is not the case, the points outside of the disc of radius ω_n centred at q do not contribute). Putting the bounds together yields

$$\mathbb{P}(\Lambda \geq x + t) \leq 8 \exp(-\omega_n^{3/2}),$$

and the claim follows by observing that for

$$c := \frac{\mathbb{E}[2(1 + \text{Po}((\pi - A)\mathcal{R}^2))\mathcal{R}]}{\mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)]} = \frac{2\mathbb{E}[\mathcal{R}] + \mathbb{E}[2(\pi - A)\mathcal{R}^3]}{\mathbb{E}[\mathcal{R}(1 + \cos 2\alpha)]},$$

it is the case that $cL_0 + (3\sqrt{L_0} + 1)\omega_n^4 \geq x + t$ for all n large enough. Simple integration using the distributions of \mathcal{R} and α then yields the expression in (11). \square

3.3 The number of sites accessed

In this section, we bound the total number of sites accessed by the cone-walk algorithm, counted with multiplicity. We note a point is accessed at step i if it is the endpoint of an edge whose other end lies inside the disc D_i . A given point may be accessed more than once, but via different edges, so we will bound the number of edges such that for some i , one end point lies inside D_i and the other outside; we call these *crossing edges*. (Note that the number of crossing edges does not quite bound the complexity of the algorithm for the complexity of a given step is not linear in the number of accessed points; however it gives very good information on the amount of data the algorithm needs to process.)

Proposition 18. *Let $A = A(z)$ be the number of sites in Φ accessed by the cone walk algorithm (with multiplicity) when walking towards q from z . Then there exists a constant $c > 0$ such that*

$$\mathbb{P}\left(A(z) > cL_0 + 4\left(\sqrt{L_0} + 1\right)\omega_n^6\right) \leq 3 \exp(-\omega_n^{5/4}).$$

For n sufficiently large.

Proof. In order to bound the number of such edges, we adapt the concept of the *border point* introduced by Bose and Devroye [5] to bound the stabbing number of a random Delaunay triangulation. For $B \subseteq \mathcal{D}_n$ and a point $x \in \mathcal{D}_n$, let $\|xB\| := \inf\{\|xy\| : y \in B\}$ denote the distance from x to B .

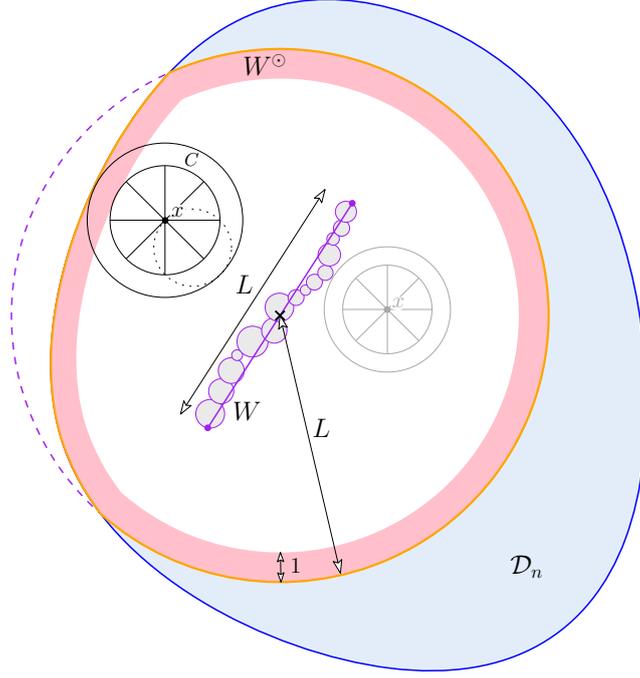


Figure 9: For the proof of Proposition 18.

We consider the walk from z to q in \mathcal{D}_n , letting

$$W = \bigcup_{i=1}^{\kappa} D_i \quad \text{and} \quad W^\circ := b(\mathbf{w}, 2 \max\{\|zq\|, \omega_n^5\}),$$

where \mathbf{w} denotes the centroid of the segment zq and we recall that $b(\mathbf{x}, r)$ denotes the closed ball centred at \mathbf{x} of radius r . Then, for $x \in W^\circ$, let C be the disc centred at x and with radius $\min\{\|xW\|, \|x\partial W^\circ\|\}$. Partition the disc C into 8 isometric cone-shaped sectors (such that one of the separation lines is vertical, say) truncated to a radius of $\sqrt{3}/2$ times that of the outer disc (see Figure 9). We say that x is a border point if one of the 8 is a border point then there is no Delaunay edge between x and a point lying outside C , since a circle through x and $y \notin C \subset W^\circ \setminus W$ must entirely enclose at least one sector of C (see dotted circle in Figure 9). Thus if x has a Delaunay edge with extremity in W , then x must be a border point. The connection between border points and the number of crossing edges can be made via Euler's relation, since it follows that a crossing edge is an edge of the (planar) subgraph of $\text{DT}(\Phi)$ induced by the points which either lie inside W , are border points, or lie outside of W° and have a neighbour in W . Let B_W denote the set of border points, E_W the collection of crossing edges, and Y_W the collection of points lying outside of W° and having a Delaunay neighbour within W . Then

$$A(z) \leq |E_W| \leq 3(|W \cap \Phi| + |B_W| + |Y_W|). \quad (15)$$

Proposition 16 bounds $|W \cap \Phi|$, as this is exactly the set of *visited vertices*. Lemmas 19 and 20 bounding $|B_W|$ and $|Y_W|$ complete the proof. \square

Lemma 19. *For all n large enough, we have*

$$\mathbb{P}\left(|Y_W| \geq 10 \max\{L_0, \omega_n^5\}\right) \leq 2 \exp\left(-\omega_n^{5/4}\right).$$

Proof. Heuristically, our proof will follow from the fact that, with high probability, a Delaunay edge away from the boundary of the domain is not long enough to span the distance between a point within the walk, and a point outside of W° . Unfortunately our proof is complicated by points on the walk which are very close to the boundary of the domain, since in this case, those points might have ‘bad’ edges which are long enough to escape W° . To deal with this, we will take all points in the walk that are close to the border, and imagine that every Delaunay edge touching one of these points is such a ‘bad’ edge. The total number of these edges will be bounded by the maximum degree.

To begin, we give the first case. Consider an arbitrary point $x \in W \cap \Phi$ that is at least ω_n^{-3} away from the boundary of \mathcal{D}_n . Suppose this point has a neighbour outside of W° , then its circumcircle implicitly overlaps an unconditioned region of \mathcal{D}_n with area at least $c\omega_n^{-3}\omega_n^5 = c\omega_n^2$ (for $c > 0$ a constant depending on the shape of the domain). The probability that this happens for x is thus at most $\exp(-\omega_n^2)$. Now note that there are at most $2n$ points in Φ with probability bounded by $\exp(-\omega_n^2)$ and at most $4n^2$ edges between points of $x \in W \cap \Phi$ and $x \in \{W^\circ\}^c \cap \Phi$. By the union bound, the probability that any such edge exists is at most

$$(4n^2) \exp(-c\omega_n^2) + \exp(-\omega_n^2) \leq \exp(-\omega_n^{3/2}).$$

For the second case, we count the number of points within ω_n^{-3} of the boundary of the domain. Using standard arguments, we have that there are no more than $10 \max\{L, \omega_n^5\} \cdot \omega_n^{-3}$ such points, with probability at least $\exp(-\omega_n^2)$. Each of these has at most Δ_Φ edges that could exit W° , where Δ_Φ is the maximum degree of any vertex in $\text{DT}(\Phi)$, which is bounded in Proposition 30. Thus, the number of such bad edges is at most $10 \max\{L, \omega_n^5\} \omega_n^{-3} \cdot \omega_n^3$ with probability at least $\exp(-\omega_n^2) + \exp(-\omega_n^{5/4})$ \square

Lemma 20. *For all n large enough, and universal constant $C > 0$,*

$$\mathbb{P}\left(|B_W| \geq C \max\{L_0, \omega_n^6\}\right) \leq 2 \exp\left(-\omega_n^{3/2}\right).$$

Proof. Since $|B_W|$ is a sum of indicator random variables, it can be bounded using (a version of) Chernoff–Hoeffding’s method. The only slight annoyance is that the indicators $\mathbf{1}_{\{x \in B_W\}}$, $x \in \Phi \cap W^\circ$ are not independent. Note however that $\mathbf{1}_{\{x \in B_W\}}$ and $\mathbf{1}_{\{y \in B_W\}}$ are only dependent if the discs used to define membership to B_W for x and y intersect. There is a priori no bound on the radius of these discs, and so we shall first discard the points $x \in \Phi$ lying far away from ∂W° and W . More precisely, let B_W^* denote the set of border points lying within distance ω_n of either W or ∂W° , and $B_W^\bullet = B_W \setminus B_W^*$. Observe now that we may bound $|B_W^\bullet|$ directly

using Lemma 10, since a point is only a border point if one of its cones is empty, and each such empty cone contains a large empty circle. So for sufficiently large n ,

$$\mathbb{P}(|B_W^\bullet| \neq 0) \leq \exp\left(-\omega_n^{3/2}\right). \quad (16)$$

Bounding $|B_W^\star|$ is now easy since the amount of dependence in the family $\mathbf{1}_{\{x \in B_W\}}, x \in \Phi \setminus W$ is controlled and we can use the inequality by Janson [15, 16]. We start by bounding the expected value $\mathbb{E}|B_W^\star|$. Note that for a single point $x \in \Phi_n$, by definition the disc used to define whether x is a border point does not intersect W and stays entirely within \mathcal{D} , so

$$\begin{aligned} \mathbb{P}_x(x \in B_W^\star) &= \mathbb{P}(x \in B_W^\star) \leq 24 \exp\left(-\frac{\pi}{32} \min\{\|xW\|, \|x\partial W^\circ\|\}^2\right) \\ &\leq 24 \exp\left(-\frac{\pi}{32} \|xW\|^2\right) + 24 \exp\left(-\frac{\pi}{32} \|x\partial W^\circ\|^2\right) \end{aligned} \quad (17)$$

and Φ_n is unconditioned in $\mathcal{D} \setminus W$. Partition $\mathcal{D} \setminus W$ into disjoint sets as follows:

$$\mathcal{D} \setminus W = \bigcup_{i=0}^{\infty} U_i$$

where $U_i := \{x \in \mathcal{D} : i \leq \|xW\| < i+1\}$. Similarly, the sets $U'_i := \{x \in W^\circ : i \leq \|x\partial W^\circ\| < i+1\}$ form a similar partition for W° . Writing λ for the 2-dimensional Lebesgue measure and using (17) above, we have

$$\begin{aligned} \mathbb{E}|B_W^\star| &= \mathbb{E} \left[\sum_{x \in \Phi} \mathbf{1}_{\{x \in W^\circ \setminus W\}} \mathbf{1}_{\{x \in B_W^\star\}} \right] \\ &= \int_{W^\circ \setminus W} \mathbb{P}_x(x \in B_W^\star) \lambda(dx) \\ &\leq \sum_{i=0}^{\infty} \int_{U_i} 24 \exp(-\pi i^2/32) \lambda(dx) + \sum_{i=0}^{\infty} \int_{U'_i} 24 \exp(-\pi i^2/32) \lambda(dx) \\ &= 24 \sum_{i=0}^{\infty} (\lambda(U_i) + \lambda(U'_i)) \exp(-\pi i^2/32). \end{aligned}$$

We may now bound $\lambda(U_i)$ and $\lambda(U'_i)$ as follows. Recall that W is a union of discs $W = \cup_i D_i$. We clearly have that

$$U_i \subseteq \bigcup_{j=0}^{\kappa-1} \{x \in \mathcal{D} : i \leq \|xD_j\| < i+1\}.$$

Note that

$$\lambda(\{x \in \mathcal{D} : i \leq \|xD_j\| < i+1\}) \leq \pi((R_j + i + 1)^2 - (R_j + i)^2) \quad (18)$$

$$= \pi(2(R_j + i) + 1). \quad (19)$$

So, assuming there are κ steps in the walk we get

$$\lambda(U_i) \leq \sum_{j=0}^{\kappa-1} \pi(2(R_j + i) + 1) = 2\pi \sum_{j=0}^{\kappa-1} R_j + \pi(i + 1)\kappa.$$

Regarding $\lambda(U'_i)$, note first that W° is convex for it is the intersection of two convex regions. It follows that its perimeter is bounded by $4\pi \max\{\|zq\|, \omega_n\}$, so that $\lambda(U'_i) \leq 4\pi \max\{\|zq\|, \omega_n\}$ for every $i \geq 0$. It now follows easily that there exist universal constants C, C' such that

$$\begin{aligned} \mathbb{E} |B_W^*| &= \mathbb{E} \mathbb{E}[|B_W^*| \mid R_i, i \geq 0] \leq C \mathbb{E} \left[\sum_{j=0}^{\kappa-1} R_j + \kappa + \max\{\|zq\|, \omega_n\} \right] \\ &\leq C' \max\{\|zq\|, \omega_n\}. \end{aligned}$$

For the concentration, we use the fact that if $\|xW\|, \|x\partial W^\circ\| \leq \omega_n$ then the chromatic number χ of the dependence graph of the family $\mathbf{1}_{\{x \in B_W^*\}}$ is bounded by the maximum number of points of Φ_n contained in a disc of radius $2\omega_n$. We then have

$$\begin{aligned} \mathbb{P}(\chi \geq 8\pi\omega_n^2) &\leq \mathbb{P}(\exists x \in \mathcal{D} : \Phi_n \cap b(x, 2\omega_n)) \\ &\leq \mathbb{E} \left[\sum_{x \in \Phi_n} \mathbb{P}_x(|b(x, 2\omega_n) \cap \Phi_n| \geq 8\pi\omega_n^2) \right] \\ &\leq \mathbb{E} \left[\sum_{x \in \Phi_n} \mathbb{P}(\text{Po}(4\pi\omega_n^2) \geq 8\pi\omega_n^2) \right] \\ &\leq \exp(-\omega_n^2), \end{aligned}$$

for all n large enough, using the bounds for Poisson random variables we have already used in the proof of Corollary 15. Let

$$W^\partial := \{x \in W^\circ \mid \max\{\|xW\|, \|x\partial W^\circ\|\} \leq \omega_n\}.$$

Following Equation (18) and by the convexity of W° , there exists a universal constant C'' such that

$$\mathbb{P}\left(|\Phi_n \cap W^\partial| \geq C'' \max\{L, \omega_n^5\} \cdot \omega_n^2\right) \leq \exp(-\omega_n^2).$$

By Theorem 3.2 of [15], we thus obtain for $t > 0$,

$$\begin{aligned} &\mathbb{P}(|B_W^*| \geq \mathbb{E} |B_W^*| + t) \\ &\leq \mathbb{E} \left[\exp\left(-\frac{2t^2}{\chi \cdot |\Phi_n \cap W^\partial|}\right) \right] \\ &\leq \exp\left(-\frac{t^2}{8\pi C'' \max\{L_0, \omega_n^5\} \cdot \omega_n^4}\right) + \mathbb{P}\left(\chi \geq 8\pi\omega_n^2\right) + \mathbb{P}\left(|\Phi_n \cap W^\partial| > C'' \max\{L_0, \omega_n^5\} \cdot \omega_n^2\right) \\ &\leq \exp\left(-\frac{t^2}{8\pi C'' \max\{L_0, \omega_n^5\} \cdot \omega_n^4}\right) + 2 \exp(-\omega_n^2). \end{aligned}$$

The result follows for n sufficiently large by choosing $t := C''(L_0 + \omega_n^6)$. \square

3.4 Algorithmic complexity

Whilst we have given explicit bounds on the number of sites in Φ_n that may be accessed by an instance of CONE-WALK, we recall that the complexity of the algorithm CONE-WALK(z, q) does not follow directly. This is because the CONE-WALK algorithm must do a small amount of computation at each step in order to compute the vertex which should be chosen next. We proceed by defining a random variable $T(z, q)$, which will denote the number of operations required by CONE-WALK(z, q) in the RAM model of computation given an implementation based upon a priority queue. We conjecture that the bound for Proposition 21 given in this section is not tight, and that the algorithmic complexity is in fact bounded by $O(L_0 + \omega_n^4)$. Unfortunately the dependency structure in algorithms of this type makes such bounds difficult to attain.

Proposition 21. *Let $T(z, q)$ be the number of steps required by the CONE-WALK algorithm to compute the sequence of stoppers given by the cone-walk process between z and q along with the path generated by SIMPLE-PATH in the RAM model of computation. Let $c_1 > 0$ be an implementation-dependent constant, then for n large enough, we have*

$$\mathbb{P}\left(T(z, q) > c_1 \cdot L_0 \log \omega_n + \omega_n^4\right) \leq 10 \exp\left(-\omega_n^{3/2}\right).$$

Proof. Define M_i to be the number of sites accessed during the i 'th step of the algorithm. We fix z, q and write T to denote $T(z, q)$ for brevity. From Section 2.3, we know that for sufficiently large n , we may choose a constant c_2 such that

$$T \leq c_2 \sum_{i=0}^{\kappa-1} \tau_i \log(M_i). \quad (20)$$

So that to bound T , it suffices to bound the sum in (20). We have

$$\begin{aligned} & \mathbb{P}\left(\sum_{i=0}^{\kappa-1} \tau_i \log(M_i) \geq 15(L_0 + \omega_n^3) \log(\omega_n)\right) \\ & \leq \mathbb{P}\left(\sum_{i=0}^{\kappa-1} \tau_i \log(M_i) \geq 3(L_0 + \omega_n^3) \log(\omega_n^5) \mid M_{\max} < \omega_n^5\right) + \mathbb{P}\left(M_{\max} \geq \omega_n^5\right) \\ & \leq \mathbb{P}\left(\sum_{i=0}^{\kappa-1} \tau_i \geq 3(L_0 + \omega_n^3)\right) + \exp\left(-\omega_n^{3/2}\right). \end{aligned}$$

The first term of which is bounded by Proposition 16, and the bound on M_{\max} comes directly from the proof of Proposition 11. \square

4 Relaxing the model and bounding the cost of the worst query

4.1 About the shape of the domain and the location of the aim

The analysis in Section 3 was provided given the assumption that q was the centre of a disc containing Φ for clarity of exposition. We now relax the assumptions on both the shape of the domain \mathcal{D} and the location of the query q . Taking \mathcal{D} to be a disc with q at its centre ensured that $\text{Disc}(z, q, r)$ was included in \mathcal{D} for $r \leq \|zq\|$ and thus the search cone and disc were always entirely contained in the domain \mathcal{D} . If we now allow q to be close to the boundary, it may be that part of the search cone goes outside \mathcal{D} .

To begin with, we leave \mathcal{D} unchanged and allow q to be any point in \mathcal{D} . Given any point $z \in \mathcal{D}$, the convexity of \mathcal{D} ensures that the line segment zq lies within \mathcal{D} . Furthermore, one of the two halves of the disc of diameter zq is included within \mathcal{D} . Thus for any $z, q \in \mathcal{D}$ and $r \in \mathbb{R}$ the portion of $\text{Cone}(z, q, r)$ (resp. $\text{Disc}(z, q, r)$) within \mathcal{D} has an area lower bounded by half of its actual area (including the portion outside \mathcal{D}). Since the distributions of all of the random variables rely on estimations for the portions of area of $\text{Cone}(z, q, r)$ or $\text{Disc}(z, q, r)$ lying inside \mathcal{D} , we have the same order of magnitude for κ, K, Λ and T , with only a degradation of the relevant constants. The proofs generalise easily, and we omit the details. (Note however, that upper and lower bounds in an equivalent of Proposition 14 would not match any longer.)

The essential property we used above is that a disc with a diameter within \mathcal{D} has one of its halves within \mathcal{D} . This is still satisfied for smooth convex domains \mathcal{D} and for discs whose radius is smaller than the minimal radius of curvature of $\partial\mathcal{D}$. Thus our analysis may be carried out provided all the cones and discs we consider are small enough. The conditioning on the event G_k which we used in Section 3 precisely guarantees that for all n large enough, on G_k , all the regions we consider are small enough ($O(\log n/\sqrt{n}) = o(1)$ in this scaling), and that G_k still occurs with high probability. These remarks yield the following result. As before, $\mathcal{D}_n = \sqrt{n}\mathcal{D}$ denotes the scaling of \mathcal{D} with area n .

Proposition 22. *Let \mathcal{D} be a fixed smooth convex domain of area 1 and diameter δ . Consider a Poisson point process Φ_z^n of intensity 1 contained in $\mathcal{D}_n = \sqrt{n}\mathcal{D}$. Let $z, q \in \mathcal{D}_n$. Let Φ_z^n be Φ^n conditioned on $z \in \Phi^n$. Then, there exist constants $C_{\Gamma, \mathcal{D}}, A_{\Gamma, \mathcal{D}}, \Gamma \in \{\kappa, K, \Lambda, T\}$, such that for the cone walk on Φ_z^n , and all n large enough, we have*

$$\sup_{z, q \in \mathcal{D}_n} \mathbb{P} \left(\Gamma(z, q) > C_{\Gamma, \mathcal{D}} \cdot \|zq\| + (1 + \sqrt{\|zq\|})\omega_n^5 \right) \leq A_{\Gamma, \mathcal{D}} \cdot \exp \left(-\omega_n^{3/2} \right).$$

Thus we obtain upper tail bounds for the number of steps $\kappa(z, q)$, the number of visited sites $K(z, q)$, the length $\Lambda(z, q)$ and the complexity $T(z, q)$ which are *uniform* in the starting point z and the location of the query q . We now move on to strengthening the results to the worst queries (still in a random Delaunay).

4.2 The worst query in a Poisson Delaunay triangulation

In this section, we prove a Poissonised version of Theorems 1 and 3. The proof of the latter are completed in Section 4.3. Consider $\sup_{z \in \Phi^n, q \in \mathcal{D}_n} \Gamma(z, q)$, the value of the parameter for the *worst possible* pair of starting point and query location, for $\Gamma \in \{\kappa, K, \Lambda, T\}$.

Theorem 23. *There exist a constant $C_{\Gamma, \mathcal{D}}$ depending only on Γ and on the shape of \mathcal{D} such that, for all n large enough,*

$$\mathbb{P}\left(\exists z \in \Phi^n, q \in \mathcal{D}_n : \Gamma(z, q) > C_{\Gamma, \mathcal{D}} \cdot \|zq\| + (1 + \sqrt{\|zq\|}) \log^4 n\right) \leq \frac{1}{n^2}.$$

By Lemma 9, the number of possible walks in a given Delaunay tessellation of size n is at most $n \times n^4$. Although it *seems* intuitively clear that this should be sufficient to bound the parameters for the worst query, it is not the case: one needs to guarantee that there is a way to sample points *independently* from Φ^n such that all the cells are hit, or in other words that the cells are not too small. We prove the following:

Proposition 24 (Stability of the walk). *There exists a partition of \mathcal{D} into at most $2n^4$ cells such that the sequence of steps used by the cone walk algorithm from any vertex of the triangulation does not change when q moves in a region of the partition. Furthermore, let $a(\Phi^n)$ denote the area of the smallest cell. Then*

$$\mathbb{P}(a(\Phi^n) < \eta) < 1 - O(n^8 \eta^{1/3}),$$

for all n large enough.

The proof of Proposition 24 relies on bounds on the smallest angle and on the shortest line segment in the line arrangement. The argument is rather long, and we think it would dilute too much the focus of this section, so we prove it in Section 5.

Proof of Theorem 23. In order to control the behaviour of the walk aiming at any point $q \in \mathcal{D}$, it suffices to control it for one point of each face of the subdivision associated with the line arrangement $\Xi(\Phi^n)$ introduced in Section 2.4. Let $\mathcal{Q}_n = \{q_i : 1 \leq i \leq n^{2k}\}$, for some $k \geq 1$ to be chosen later, be a collection of i.i.d. uniform random points in \mathcal{D} , independent of Φ^n . Let E_n be the event that every single face of the subdivision contains at least one point of \mathcal{Q}_n . Since there are at most $|\Phi^n|^4$ regions, Proposition 24 implies that

$$\begin{aligned} \mathbb{P}(E_n^c) &\leq \mathbb{P}(E_n^c \mid a(\Phi^n) \geq n^{-k}) + \mathbb{P}(a(\Phi^n) < n^{-k}) \\ &\leq \mathbb{E}[|\Phi^n|^4] \cdot \mathbb{P}(\text{Bin}(n^{2k}, n^{-k}) = 0) + O(n^8 \cdot n^{-k/3}) \\ &= O(n^4 \cdot \exp(-n^{k/2})) + O(n^{8-k/3}) \\ &= O(n^{8-k/3}). \end{aligned} \tag{21}$$

For $x > 0$ write $f_n(x) := C_{\Gamma, \mathcal{D}}x + (1 + \sqrt{x}) \log^4 n$. Then, we have

$$\begin{aligned} &\mathbb{P}(\exists z \in \Phi^n, q \in \mathcal{D}_n : \Gamma(z, q) \geq f_n(\|zq\|)) \\ &\leq \mathbb{P}(\exists z \in \Phi^n, q \in \mathcal{D}_n : \Gamma(z, q) \geq f_n(\|zq\|) \mid E_n) + \mathbb{P}(E_n^c) \\ &= \mathbb{P}(\exists z \in \Phi^n, q \in \mathcal{Q}_n : \Gamma(z, q) \geq f_n(\|zq\|)) + \mathbb{P}(E_n^c) \\ &\leq n^{2k} \cdot \sup_{q \in \mathcal{D}} \mathbb{P}(\exists z \in \Phi^n : \Gamma(z, q) \geq f_n(\|zq\|)) + \mathbb{P}(E_n^c). \end{aligned} \tag{22}$$

Now for any fixed $q \in \mathcal{D}$, and conditioning on Φ^n , we see that

$$\begin{aligned} \mathbb{P}(\exists z \in \Phi^n : \Gamma(z, q) \geq f_n(\|zq\|)) &\leq \mathbb{E} \left[\sum_{z \in \Phi^n} \mathbf{1}_{\{\Gamma(z, q) \geq f_n(\|zq\|)\}} \right] \\ &\leq \mathbb{E} \left[\sum_{z \in \Phi^n} \mathbf{1}_{\{\Gamma(z, q) \geq f_n(\|zq\|), |\Phi^n| \leq 2n\}} \right] + \mathbb{P}(|\Phi^n| > 2n) \\ &\leq 2n \sup_{z \in \mathcal{D}} \mathbb{P}(\Gamma(z, q) \geq f_n(\|zq\|)) + \exp(-n/3). \end{aligned}$$

The claim follows from (21), (22), and Proposition 22 by choosing $k = 40$. (Note that we could easily obtain sub-polynomial bounds.) \square

Before concluding this section, we note that the same arguments also yield the Poissonised version of Theorem 1 about the number of neighbourhoods required to compute all the steps of any query. We omit the proof.

Theorem 25. *We have the following bound for any run of CONE-WALK on $\text{DT}(\Phi^n)$, for Φ^n a Poisson point process of rate n in \mathcal{D} :*

$$\mathbb{P}(\exists z \in \Phi^n, q \in \mathcal{D}_n : M(z, q) > \log^{1+\epsilon} n) \leq \frac{1}{n^2}.$$

4.3 De-Poissonisation: Proof of Theorem 3

In this section, we prove that the results proved in the case of a Poisson point process of intensity n in a compact convex domain \mathcal{D} (Theorem 23) may be transferred to the situation where the collection of sites consists of n independent uniformly random points in \mathcal{D} . In the present case, the concentration for our events is so strong that the de-Poissonisation is straightforward.

Note that conditional on $|\Phi^n| = n$, the collection of points Φ^n is precisely distributed like n independent uniforms in \mathcal{D} . Furthermore, $|\Phi^n|$ is a Poisson random variable with mean n , so that by Stirling's formula, as $n \rightarrow \infty$,

$$\mathbb{P}(|\Phi^n| = n) = \frac{n^n \exp(-n)}{n!} \sim \frac{1}{\sqrt{2\pi n}}. \quad (23)$$

This is small, but we can consider multiple copies of the process; if one happens to have exactly n points, but that none of them actually behaves badly, then it must be the case that the copy that has n points behave as it should, which is precisely what we want to prove. To make this formal, consider a sequence $\Phi_i^n, i \geq 1$, of i.i.d. Poisson point processes with mean n in \mathcal{D} . Let η_n be the first $i \geq 1$ for which $|\Phi_i^n| = n$. Then, for any event E^i defined on Φ_i^n ,

$$\begin{aligned} \mathbb{P}(E_1 \mid |\Phi_1^n| = n) &= \mathbb{P}(E_{\eta_n}) \\ &\leq \mathbb{P}(\exists j \leq \eta_n : E_j) \\ &\leq \mathbb{P}(\exists j \leq n : E_j, \eta_n \leq n) \\ &\leq n\mathbb{P}(E_1) + \mathbb{P}(\eta_n \geq n) \\ &\leq n\mathbb{P}(E_1) + \exp\left(-\sqrt{n/(2\pi)}\right), \end{aligned}$$

where the last line follows from (23). In the present case, we apply the argument above to estimate the probabilities that the number of steps, the number of sites visited, the length of the walk or the complexity of the algorithm exceeds a certain value. In the case of the Poisson point process, we have proved that any of these events has probability at most $1/n^2$, so that we immediately obtain a bound of $O(1/n)$ in the case of n i.i.d. uniformly random points, which proves the main statement in Theorem 3.

This implies immediately that, for δ the diameter of \mathcal{D} , we have

$$\begin{aligned} \mathbb{E} \left[\sup_{z \in \mathbf{X}^n, q \in \mathcal{D}_n} \Gamma(z, q) \right] &\leq 2\delta C_\Gamma \sqrt{n} + \mathbb{P} \left(\sup_{z \in \mathbf{X}^n, q \in \mathcal{D}_n} \Gamma(z, q) \geq 2\delta C_\Gamma \sqrt{n} \right) \\ &\leq 2\delta C_\Gamma \sqrt{n} + O(1/n) \\ &\leq 3\delta C_\Gamma \sqrt{n}, \end{aligned}$$

for n large enough, so that the proof of Theorem 3 is complete. The proof of Theorem 1 from Theorem 25 relies on the same ideas and we omit the proof.

5 Extremes in the arrangement $\Xi(\Phi^n)$

In this section, we obtain the necessary results about the geometry of the line arrangement Ξ introduced in Section 2.4.4. In particular, we prove Proposition 24 providing a uniform lower bound on the areas of the cells of the subdivision associated to Ξ which is crucial to the proofs of our main results Theorems 1 and 3.

The rays (half-lines) of the arrangement Ξ all have one end at a site of Φ^n . Let Ξ_x denotes the set of rays of Ξ with one end at x ; we say that the rays in Ξ_x originate at x . The rays come in two kinds:

- *regular rays* are defined by a triple x, y, y' such that there exists $r > 0$ and $q \in \mathbb{R}^2$ for which y and y' both lie on the front arc of the $\text{Cone}(x, q, r)$; such a ray is denoted by $\rho^-(x, y, y')$;
- *extreme rays* correspond to one of the straight boundaries of $\text{Cone}(x, y, \infty)$, for some pair of points x, y ; these two rays are denoted by $\rho_1^+(x, y)$ and $\rho_2^+(x, y)$ in such a way that the angle from ρ_1^+ to ρ_2^+ is $\pi/4$.

We let Ξ_x^- and Ξ_x^+ denote the collections of regular and extreme rays of Ξ_x , respectively. Our strategy to bound the area of the smallest cell is to lower bound the angle between any two rays (Section 5.1) and the length of any line segment (Section 5.2). We put together the arguments and prove Proposition 24 in Section 5.3.

5.1 The smallest angle between two rays

We start with a bound on the smallest angle in the arrangement.

Lemma 26. *Let $\langle \Phi^n \rangle$ denote the minimum angle between any two lines of $\Xi(\Phi^n)$ which intersect within \mathcal{D} . Then, there exists a constant C such that, for any $\beta \in (0, \pi/8)$,*

$$\mathbb{P}(\langle \Phi^n \rangle < \beta) \leq Cn^5 \beta.$$

Proof. For two intersecting rays ρ_1 and ρ_2 , let $\langle \rho_1, \rho_2 \rangle$ denote the (smallest) angle they define. We first deal with the angles between two rays originating from different points. Consider a given ray $\rho_1 \in \Xi_{x_1}$, and let \tilde{x}_1 denote the intersection of ρ_1 with $\partial\mathcal{D}$. For another ray $\rho_2 \in \Xi_{x_2}$, with $x_2 \neq x_1$, to intersect ρ_1 at an angle lying in $(-\beta, \beta)$, we must have $x_2 \in \text{Cone}_\beta(x_1, \tilde{x}_1, \infty) \cup \text{Cone}_\beta(\tilde{x}_1, x_1, \infty)$. Observe that for any two points x, y , the area of $\text{Cone}_\beta(x, y, \infty) \cap \mathcal{D}$ is bounded by $\delta^2 \tan \beta$, where δ denotes the diameter of \mathcal{D} . In particular, for any fixed ray $\rho_1 \in \Xi_x$ for some $x \in \Phi^n$,

$$\begin{aligned} & \mathbb{P}(\exists x_2 \in \Phi_x^n \setminus \{x\}, \rho_2 \in \Xi_{x_2} : \langle \rho_1, \rho_2 \rangle < \beta) \\ & \leq \mathbb{P}(\exists x_2 \in \Phi_x^n \setminus \{x\} : x_2 \in \text{Cone}_\beta(x_1, \tilde{x}_1, \infty)) \\ & \quad + \mathbb{P}(\exists x_2 \in \Phi_x^n \setminus \{x\} : x_2 \in \text{Cone}_\beta(\tilde{x}_1, x_1, \infty)) \\ & \leq 2n\delta^2 \tan \beta, \end{aligned}$$

since $\Phi_x^n \setminus \{x\}$ is a Poisson point process in \mathcal{D} with intensity n . Now, since for any x , $|\Xi_x| \leq 2(|\Phi^n| - 1)$,

$$\begin{aligned} & \mathbb{P}(\exists x_1, x_2 \in \Phi^n, x_1 \neq x_2, \rho_1 \in \Xi_{x_1}, \rho_2 \in \Xi_{x_2} : \langle \rho_1, \rho_2 \rangle < \beta) \\ & \leq \mathbb{E} \left[\sum_{x \in \Phi^n} \mathbb{P}(\exists \rho_1 \in \Xi_x, x_2 \in \Phi_x^n \setminus \{x\}, \rho_2 \in \Xi_{x_2} : \langle \rho_1, \rho_2 \rangle < \beta) \right] \\ & \leq \mathbb{E} \left[\sum_{x \in \Phi^n} 2(|\Phi^n| - 1) \cdot 2n\delta^2 \tan \beta \right] \\ & = 4n^3 \delta^2 \tan \beta, \end{aligned} \tag{24}$$

since $\Phi_x^n \setminus \{x\}$ is distributed like Φ^n .

We now deal with the smallest angle between any two rays in the same set Ξ_x , for some $x \in \Phi^n$. Any point $y \in \Phi_x^n \setminus \{x\}$ defines at most two extreme rays in Ξ_x which intersect at an angle of $\pi/4$. For any two distinct points y, y' to define two extreme rays that intersect at an angle smaller than $\beta > 0$, then y' must lie in $\text{Cone}_\beta(x, y, \infty)$ or one of its two images in the rotations of angle $+\pi/8$ or $-\pi/8$ about x . The arguments we have used to obtain (24) then imply that

$$\begin{aligned} \mathbb{P}(\exists x \in \Phi^n : \exists \rho, \rho' \in \Xi_x^+, \langle \rho, \rho' \rangle < \beta) & \leq 3n^3 \delta^2 \tan \beta \\ & \leq 6n^3 \delta^2 \beta, \end{aligned}$$

for all $\beta \in [0, \pi/4]$, since then $\tan \beta \leq 2\beta$.

We now consider a regular ray $\rho = \rho^-(x, y, z)$. Then there exist $q \in \mathbb{R}^2$ and $r > 0$ such that ρ is the half-line with an end at x and going through q , and x, y, z all lie on the same circle \mathcal{C} (more precisely, y, z lie on the intersection of \mathcal{C} with $\text{Cone}(x, q, \infty)$). We first bound the probability that there exists a point z' such that the regular ray $\rho' = \rho^-(x, y, z')$ intersects ρ at an angle at most β , that is $\rho' \in \text{Cone}_\beta(x, q, \infty)$. We want to bound the area of the region

$$\{z' \in \mathcal{D} : \rho^-(x, y, z') \in \text{Cone}_\beta(x, q, \infty)\},$$

for which the angle between ρ and ρ' is at most β . By definition of ρ' , x , y and z' lie on the same circle $\mathcal{C}' = \mathcal{C}'(c)$, which has a centre c' in $\text{Cone}_\beta(x, q, \infty)$. More precisely, the centre c' lies on the intersection of the bisector of the line segment $[x, y]$ with $\text{Cone}_\beta(x, q, \infty)$; so c' lies in a line segment ℓ of length $O(\beta)$ containing the centre c of \mathcal{C} . The Hausdorff distance between \mathcal{C} and $\cup_{c' \in \ell} \mathcal{C}'(c')$ is $O(\beta)$ and it follows that

$$\mathcal{A}\{z' \in \mathcal{D} : \rho^-(x, y, z') \in \text{Cone}_\beta(x, q, \infty)\} = O(\beta).$$

A similar argument applies to deal with rays $\rho^-(x, y', z')$ for $y', z' \in \Phi^n$: once the one of y' or z' is chosen, the second is forced to lie in a region of area $O(\beta)$. From there, arguments similar to the ones we used in (24) yield

$$\mathbb{P}(\exists x \in \Phi^n : \exists \rho, \rho' \in \Xi_x^-, \langle \rho, \rho' \rangle < \beta) = O(n^5 \beta). \quad (25)$$

Finally, we deal with the intersections of a regular and an extreme ray. We $\rho = \rho^-(x, y, z)$ and $\rho' = \rho_1^+(x, y')$, for $y' \in \Phi^n$ (potentiall $y' = y$). For ρ' to lie in $\text{Cone}_\beta(x, q, \infty)$, the point y' must lie in a cone with apex x , half-angle β and with axis one of the straight boundaries of $\text{Cone}(x, q, \infty)$. So if $y \neq y'$, the point y' is constrained to lie in a region of area $O(\beta)$; in this case, the expected number of $x, y, z, y' \in \Phi^n$ in such a configuration is $O(n^4 \beta)$. If on the other hand, $y' = y$ then y must lie on a portion of length $O(\beta)$ of the front arc of $\mathcal{C} \cap \text{Cone}(x, q, \infty)$. However, conditional on x, y, z indeed forming the ray $\rho = \rho^-(x, y, z)$, the points y is uniformly distributed on the front arc. In other words, the expected number of triplets x, y, z in such a configuration is $O(n^3 \beta)$. It follows that

$$\mathbb{P}(\exists x \in \Phi^n : \exists \rho \in \Xi_x^-, \rho' \in \mathcal{S}_x^+, \langle \rho, \rho' \rangle < \beta) = O(n^4 \beta). \quad (26)$$

Putting (24), (25) and (26) together completes the proof. \square

5.2 The shortest line segment

Let $L(\Phi^n)$ denote the smallest line segment in the arrangement of lines $\Xi(\Phi^n)$. The aim of this section is to prove the following:

Lemma 27. *We have, for any γ small enough and n large enough,*

$$\mathbb{P}(L(\Phi^n) < \gamma) \leq n^9 \gamma^{1/2}.$$

In order to prove Lemma 27, we place a ball of radius γ at every intersection in $\Xi(\Phi^n)$ (including the ones which involve the boundary $\partial\mathcal{D}$) and estimate the probability that any such ball is hit by another ray of the arrangement. The proof is longer and slightly more delicate than that of Lemma 26 simply because three rays, each involving up to three points of Φ^n , might be involved the event of interest. The first step towards proving Lemma 27 is to establish uniform bounds on the area of location of some points for the rays to intersect a given ball. These bounds are stated in Lemmas 28 and 29 below.

Lemma 28. *There exists a constant C such that, for any $x, y, z \in \mathcal{D}$, one has, for all $\gamma > 0$ small enough, we have*

- (i) $\mathcal{A}\{\bar{x} \in \mathcal{D} : \rho^+(\bar{x}, y) \cap \delta(z, \gamma) \neq \emptyset\} \leq C\gamma$
- (ii) $\mathcal{A}\{\bar{y} \in \mathcal{D} : \rho^+(x, \bar{y}) \cap \delta(z, \gamma) \neq \emptyset\} \leq C\gamma/\|xz\|$.

Proof. Fix x, z and $\gamma > 0$. The location of points \bar{y} for which $\rho^+(x, \bar{y}) \ni z$ is precisely $\partial \text{Cone}(x, z, \infty)$. Then,

$$\bigcup_{q \in \delta(z, \gamma)} \partial \text{Cone}(x, q, \infty)$$

consists of two cones of half-angle $\arcsin(\gamma/\|xz\|)$, which proves the second claim.

For the first claim, let $c_1 = c_1(z)$ and $c_2 = c_2(z)$ be the two intersections of the circles of radius $\|zy\|/\sqrt{2}$ centred at y and z (so the segment zy is seen from c_1 and c_2 at an angle of $\pi/2$). Let $\mathcal{C}_1 = \mathcal{C}_1(z)$ and $\mathcal{C}_2 = \mathcal{C}_2(z)$ denote the discs centred at c_1 and c_2 , respectively, and having y and z on their boundaries. Then, the region of the points \bar{x} for which $\rho^+(\bar{x}, y) \ni z$ is precisely the boundary of $\mathcal{C}_1 \cup \mathcal{C}_2$. When looking for the region of points \bar{x} satisfying $\rho^+(\bar{x}, y) \cap \delta(z, \gamma) \ni q$, for some $q \in \delta(z, \gamma)$ observe that the centres $\|c_1(q) - c_1(z)\| = O(\gamma)$, $\|c_2(q) - c_2(z)\| = O(\gamma)$ and that the Hausdorff distance between $\mathcal{C}_1(q) \cup \mathcal{C}_2(q)$ and $\mathcal{C}_1 \cup \mathcal{C}_2$ is $O(\gamma)$, hence the first claim. \square

Lemma 29. *There exists a constant C such that, for any $x, y, y', z \in \mathcal{D}$, one has, for all $\gamma > 0$ small enough*

- (i) $\mathcal{A}\{\bar{x} \in \mathcal{D} : \rho^-(\bar{x}, y, y') \cap \delta(z, \gamma) \neq \emptyset\} \leq C\gamma$
- (ii) $\mathcal{A}\{\bar{y} \in \mathcal{D} : \rho^-(x, \bar{y}, y') \cap \delta(z, \gamma) \neq \emptyset\} \leq C\gamma$.

Proof. The ray $\rho^-(x, y, y')$ is the half-line with one end at x and going through the centre c of the circle which contains all three x, y and y' . So the region of the points \bar{x} for which $\rho^-(\bar{x}, y, y') \ni z$ is naturally indexed by the points of the bisector Δ of the segment $[y, y']$: for a point $c \in \Delta$, the only possible points $\bar{x} = \bar{x}(c)$ lie the intersection of the circle \mathcal{C} centred at c and containing y (and y') and the line containing c and z . So if $z \neq c$, there are two such points and at most one of them is such that $y, y' \in \text{Cone}(\bar{x}, c, \infty)$. If $z = c$, then there is potentially a continuum of points $\bar{x}(c)$, consisting of the arc of \mathcal{C} for which $y, y' \in \text{Cone}(\bar{x}, c, \infty)$. It follows that for every $c \in \Delta$ such that $\|zc\| > 2\gamma$, the points $\bar{x}(c)$ for which $\rho^-(\bar{x}, y, y')$ intersects $\delta(z, \gamma)$ is an arc of length $O(\gamma)$. On the other hand, the cumulated area of the circles $\mathcal{C}(c)$ for which $\|zc\| \leq 2\gamma$ is $O(\gamma)$. Hence the first claim.

For x, y and q fixed, the set of points \bar{y} such that $\rho^-(x, y, \bar{y}) \ni q$ is contained in the circle $\mathcal{C}'(q)$ whose centre is the intersection of the bisector of $[x, y]$ and the line (xq) , and which contains x (and y). The cumulated area of $\mathcal{C}'(q)$ when $q \in \delta(z, \gamma)$ is $O(\gamma)$, which proves the second claim. \square

Proof of Lemma 27. Recall that, in order to lower bound the length of the shortest line segment in $\Xi(\Phi^n)$, we prove using Lemmas 28 and 29, that if we place a ball of radius γ at every intersection in $\Xi(\Phi^n)$, every ball contains a single intersection with probability at least $1 - O(n^9\gamma^{1/2})$. The intersections to consider are of three different types: the points $x \in \Phi^n$, the intersections of one ray and the boundary of the domain \mathcal{D} , and intersections between two rays.

The case of balls centred at points of Φ^n first is easily treated. We have

$$\mathbb{P}(\exists z, x \in \Phi^n : z \neq x, \Xi_x \cap \delta(z, \gamma) \neq \emptyset) \leq \mathbb{E} \left[\sum_{z \in \Phi^n} \mathbf{1}_{\{\exists x \in \Phi^n \setminus \{z\} : \Xi_x \cap \delta(z, \gamma) \neq \emptyset\}} \right].$$

Note that, in the expected value on the right, although $x \neq z$, the set Ξ_x does depend on z in a non-trivial way. The main point about the proof is to check that the conditioning on $z \in \Phi^n$ does not make it much easier for a ray to intersect $\delta(z, \gamma)$; and this is where Lemmas 28 and 29 enter the game. We have

$$\begin{aligned} \mathbb{E} \left[\sum_{z \in \Phi^n} \mathbf{1}_{\{\exists x \in \Phi^n \setminus \{z\} : \Xi_x \cap \delta(z, \gamma) \neq \emptyset\}} \right] &\leq \mathbb{E} \left[\sum_{z \in \Phi^n} \sum_{y \in \Phi^n} \sum_{x \in \Phi^n \setminus \{z, y\}} \mathbf{1}_{\{\rho^+(x, y) \cap \delta(z, \gamma) \neq \emptyset\}} \right] \\ &+ \mathbb{E} \left[\sum_{z \in \Phi^n} \sum_{y, y' \in \Phi^n} \sum_{x \in \Phi^n \setminus \{z, y, y'\}} \mathbf{1}_{\{\rho^-(x, y, y') \cap \delta(z, \gamma) \neq \emptyset\}} \right] \\ &\leq Cn^4\gamma, \end{aligned}$$

by Lemmas 28 and 29, since conditional on $z, y, y' \in \Phi^n$, $\Phi^n \setminus \{x, y, y'\}$ is distributed like Φ^n .

We now move on to the intersections of rays with the boundary of \mathcal{D} . Let ρ be a ray, and let z denote its intersection with $\partial\mathcal{D}$. If ρ is an extreme ray, then there exists $x, y \in \Phi^n$ such that $\rho = \rho_1^+(x, y)$ (or $\rho = \rho_2^+(x, y)$) and we are interested in the probability that some other ray ρ' intersects $\delta(z, \gamma)$; note that the ray ρ' might arise from a set of points x', y' (if it is extreme) or x', y', y'' (if it is regular) which uses some of x or y . If there is at least one point of x', y' that is not in $\{x, y\}$ then Lemma 28 allows us to bound the probability that $\rho^+(x', y')$ intersects $\delta(z, \gamma)$; in the other case, there must be one of x', y', y'' which is not in $\{x, y\}$ and Lemma 29 applies: For each such choice of points, the probability that ρ' intersects $\delta(z, \gamma)$ is $O(\gamma)$, uniformly in x, y and z . So the only case remaining to check is when ρ' is only defined by the two points x, y . Here,

- either $\rho' \in \Xi_x$ and for ρ' to intersect $\delta(z, \gamma)$ one must have $x \in \delta(z, \gamma)$, which only happens if x lies within distance γ of $\partial\mathcal{D}$, hence with probability at most $O(n\gamma)$;
- or $\rho' \in \Xi_y$ and the ball $\delta(z, \gamma)$ should be close to one of the two points in $\partial\text{Cone}(x, y, \infty) \cap \partial\text{Cone}(y, x, \infty)$; Write $i_1 = i_1(x, y)$ and $i_2 = i_2(x, y)$ for these two points. More precisely, since the angle between ρ and ρ' is $\pi/4$, the point z should lie within distance $\gamma\sqrt{2}$ of one of the two intersections (or $\delta(z, \gamma)$ does not intersect ρ'). Dually, for a point q , let $\bar{y} = \bar{y}(q)$ be the point such that $i_1(x, \bar{y}) = q$. Then, for $\|zq\| \leq \gamma' = \gamma\sqrt{2}$ the angle between the vectors $x\bar{y}(q)$ and $x\bar{y}(z)$ is at most $\arcsin(\gamma'/\|xz\|)$ and the difference in length is an additive term of at most $\gamma'\sqrt{2}$. Overall, $\{\bar{y} : i_1(x, \bar{y}) \in \delta(z, \gamma')\}$ is contained in a ball of radius at most $3\gamma'\sqrt{2} = 6\gamma$. Finally, since $z \in \partial\mathcal{D}$, the probability that such a situation occurs for some $x, y \in \Phi^n$ is at most $O(n^2\gamma)$.

As a consequence, the probability that there exist two rays ρ and ρ' , such that $\rho' \cap \delta(z, \gamma) \neq \emptyset$, for $z = \rho \cap \mathcal{D}$ is $O(n^5\gamma)$.

It now remains to deal with the case of balls centred at the intersection of two rays. Consider two rays, $\rho_1 \in \Xi_{x_1}$ and $\rho_2 \in \Xi_{x_2}$, for $x_1 \neq x_2$, which are supposed to intersect at a point z (the case $x_1 = x_2$ has already been covered since then, we have $z \in \Phi^n$). The definition of ρ_1 and ρ_2 may involve up to six points x_1, y_1, y'_1 and x_2, y_2, y'_2 of Φ^n . The probability that some third ray ρ_3 , whose definition uses at least one new point of Φ^n , intersects $\delta(z, \gamma)$ can be upper bounded using Lemmas 28 and 29 as before and we omit the details. The only new cases we need to cover are the ones when the definition of ρ_3 only involves points among x_1, y_1, y'_1 and x_2, y_2, y'_2 . We now show that all the possible configurations on six points only are unlikely to occur for γ small in the random point set Φ^n .

Because of the number of cases to be treated, without a clear way to develop a big picture, we only sketch the remainder of the proof. We treat the cases where $\rho_1 = \rho^-(x_1, y_1, y'_1)$ and $\rho_2 = \rho^-(x_2, y_2, y'_2)$, so in particular the two rays which intersect at z are regular rays.

- If $\rho_3 \in \Xi_{x_1}$ (or, by symmetry, $\rho_3 \in \Xi_{x_2}$), the bound on the smallest angle between any two rays in Lemma 26 ensures that z ought to be close to x_1 : one must have $\arcsin(\gamma/\|x_1 z\|) \leq \eta$ where η denotes the smallest angle. Then, either x_2 is also close to z , or y_2 and y'_2 actually almost lie on a circle whose centre lies on the line (x_1, x_2) . Both are unlikely: (1) the closest pair of points lie at distance $\Omega(1/n)$ (which is much larger than what we are aiming for) and (2) for fixed x_1, x_2 and y_2 , the location of points y'_2 such that y_2 lies near the circle whose centre lies on (x_1, x_2) and which contains y_2 has area of order $O(\gamma)$.
- If $\rho_3 \in \Xi_{y_1}$ (or the other symmetric cases), there are the following possibilities: either ρ_3 is a regular ray and (up to symmetry)

$$\rho_3 \in \left\{ \begin{array}{ll} \text{(a)} & \rho^-(y_1, x_1, y'_1), \quad \text{(b)} & \rho^-(y_1, x_1, y_2) \\ \text{(c)} & \rho^-(y_1, x_2, y'_1), \quad \text{(d)} & \rho^-(y_1, x_2, y_2) \\ \text{(e)} & \rho^-(y_1, y'_1, y'_2), \quad \text{(f)} & \rho^-(y_1, y_2, y'_2) \end{array} \right\},$$

or ρ_3 is an extreme ray and

$$\rho_3 \in \left\{ \text{(g)} \quad \rho^+(y_1, y'_1), \quad \text{(h)} \quad \rho^+(y_1, y_2) \right\}.$$

Since $\rho^-(x_1, y_1, y'_1)$ is a ray, y_1, y'_1 both lie in a cone of half-angle $\pi/8$, and it is impossible that we also have x_1, y'_1 both lying in some cone of half-angle $\pi/8$; so configuration (a) does not occur. For situation (b), we use Lemma 29 (ii). For situations (c)–(f), we use the arguments in the proof of Lemma 29 to exhibit the constraints about the location of the third point defining ρ_3 , once the first two are chosen.

To deal with the cases where $\rho_3 \in \Xi_{y_1}^+$, observe that once y_1 is placed, y'_1 (or y_2) must lie in a cone of angle $\arcsin(\gamma/\|y_1 z\|)$: so either $\|y_1 z\| \geq r$, and the cone has area $O(\gamma/r)$ or $\|y_1 z\| \leq r$, but then y_1 must lie in the portion of the front arc of length $O(r)$ about ρ_1 . Since the conditional distribution of y_1 is uniform on the arc, the probability that this configuration occurs is $O(\inf_{r \geq 0} \{\gamma/r + r\}) = O(\gamma^{1/2})$.

The remaining cases, where at least one of ρ_1 and ρ_2 is an extreme ray, may all be treated using similar arguments and we omit the tedious details. \square

5.3 Lower bounding the area of the smallest cell: Proof of Proposition 24

With Lemmas 26 and 27 under our belt, we are now ready to prove Proposition 24.

We use the line arrangement $\Xi(\Phi^n)$ of Section 2.4. The faces of the corresponding subdivision are convex regions delimited by the line segments between intersections of rays, and pieces of the boundary of the domain \mathcal{D} . Let $L(\Phi^n)$ and $\langle\Phi^n\rangle$ denote the length of the shortest line segment and the smallest angle in the arrangement arising from Φ^n . Consider a face of the subdivision, and one of its vertices u which does not lie on the $\partial\mathcal{D}$. Then the triangle formed by the u and its two adjacent vertices on the boundary of the face is contained in the face. The lengths of the two line segments adjacent to u are at least $L(\Phi^n)$ long, and the angle they make lies between $\langle\Phi^n\rangle$ and $\pi - \langle\Phi^n\rangle$, so that the area of the corresponding triangle is at least

$$L(\Phi^n)^2 \cdot \frac{\tan\langle\Phi^n\rangle}{2}.$$

It follows easily from Lemmas 26 and 27 that the area of the smallest face is at least $\beta\gamma/2$ with probability at least $1 - O(n^4\beta) - O(n^9\gamma^{1/2})$. Choosing β, γ such that $n^4\beta = n^9\gamma^{1/2}$ yields the claim.

6 Comparison with Simulations

We implemented CONE-WALK in C++ using the CGAL libraries. For simulation purposes, we generated 10^7 points uniformly at random in a disc of area 10^7 . We then simulated CONE-WALK on 10^6 different walks starting from a point within a disc having a quarter of the radius of the outer disc (to help reduce border effects) with a uniformly random destination. For comparison, we give the expected bounds for a walk whose destination is at infinity. See Table 1. With respect to the path, we give two values for the number of extra vertices visited in a step. The first is the number obtained by using SIMPLE-PATH and the second, in brackets, is the average number of ‘intermediary vertices’ within each disc.

	Theory	Theory (5 s.f.)	Simulation
Radius	$\mathbb{E}[R] = \sqrt{\frac{\pi}{2\sqrt{2}+\pi}}$	0.72542	0.72557
# Intermediary path steps	$\leq \mathbb{E}[\tau_i] \leq \frac{4\pi}{\pi+2\sqrt{2}}$	≤ 1.1049	0.41244 (1.09574)
SIMPLE-PATH Length	$\leq \mathbb{E}[\Lambda/L_0] \leq \frac{22\pi-4\sqrt{2}}{2+3\pi+8\sqrt{2}}$	≤ 2.7907	1.51877

Table 1: Comparison of theory with simulations. Inequalities are used to show when values are bounds.

Appendices

A Maximum degree for Poisson Delaunay in a smooth convex

Let Γ be a homogeneous Poisson process on the entire Euclidean plane with intensity 1. In this case, Bern et al. [2] give a proof that the expected maximum degree of any vertex of the Delaunay triangulation $\text{DT}(\Gamma)$ falling within the box $[0, \sqrt{n}]^2$ is $\Theta\left(\frac{\log n}{\log \log n}\right)$. Whilst such a bound may be useful in the analysis of geometric algorithms, it has a shortcoming in that it implicitly avoids dealing with the border effects that occur when considering points distributed within bounded regions. When considering such bounded regions, it can be observed that the degree distribution is significantly skewed near the border, with the majority of the vertices on or near the convex hull having a much higher degree than the global average. It is therefore not altogether trivial that the maximum degree should still be bounded polylogarithmically when the sites are generated by a homogeneous Poisson process in a bounded region. In this section, we show that this is indeed the case for the specific case of a smooth compact convex. As far as we are aware, this is the first such bound that has been given for a set of random points in a bounded region. Let \mathcal{D} be a smooth compact convex subset of \mathbb{R}^2 with area $\lambda_2(\mathcal{D}) = 1$, and $\mathcal{D}_n = \sqrt{n}\mathcal{D}$ (with area n and $\text{diameter}(\mathcal{D}_n) \leq c_2\sqrt{n}$, for some constant c_2). Let Φ be a homogeneous Poisson point process of intensity 1 contained within \mathcal{D}_n , so that in expectation we have $\mathbb{E}|\Phi| = n$. We use $\delta_\Phi(x)$ to denote the degree of $x \in \Phi$ in $\text{DT}(\Phi)$, and take $\Delta_\Phi := \max_{x \in \Phi} \delta_\Phi(x)$.

Proposition 30. *For any $\xi > 0$, we have for n sufficiently large,*

$$\mathbb{P}\left(\Delta_\Phi \geq \log^{2+\xi} n\right) \leq \exp\left(-\log^{1+\xi/4} n\right).$$

Define $\|Ax\| := \inf\{\|xy\| : y \in A\}$. Our proof will follow by considering two cases. For the first case we consider all points $x \in \Phi$ satisfying $\|\partial\mathcal{D}_n x\| \leq \sqrt{\log n}$ and bound the number of neighbours of x in $\text{DT}(\Phi)$ to one side of x ; doubling the result and the end for the final bound. To begin, we trace a ray from x to the point $y \in \partial\mathcal{D}_n$ minimising $\|xy\|$; we refer to this as \mathcal{R}_0 . Next, we create a new ray, \mathcal{R}_1 exiting x such that the area enclosed by \mathcal{R}_0 , \mathcal{R}_1 and $\partial\mathcal{D}_n$ is $\log^{1+\xi} n$; we let S_0 denote this region. (See Figure 10.) For n large enough, the angle between the rays \mathcal{R}_0 and \mathcal{R}_1 is smaller than $\frac{\pi}{2} + \frac{\pi}{12}$. In addition, the length of \mathcal{R}_1 is upper bounded by the diameter of $\mathcal{D}_n \leq c_2\sqrt{n}$ and lower bounded by $c_1\sqrt{\log n}$ for $c_1 = \frac{1}{2}$ since \mathcal{R}_0 is not longer than $\sqrt{\log n}$ by hypothesis.

We then define an iterative process that finishes as soon as one of our regions (that we define shortly) is totally contained within a ball of radius $\log^{1+\xi} n$ about the point x . At each step i , we look at the ray \mathcal{R}_{i-1} and then grow a sector of a circle with area $\log^{1+\xi} n$, and of radius $\frac{1}{\sqrt{2}}|\mathcal{R}_{i-1}|$, where $|\mathcal{R}_i|$ is the length of the i th ray. The sector of radius $|\mathcal{R}_{i+1}|$ defined by the rays \mathcal{R}_i and \mathcal{R}_{i+1} is denoted by S_i . Thus the internal angle of each new sector is exactly twice that of the previous one. For each sector apart from the first, we add a ‘border’ (shaded blue in Figure 10) which extends the each sector to the length of the sector proceeding it: let Q_i be the cone delimited by the rays \mathcal{R}_i and \mathcal{R}_{i+1} , and of radius $|\mathcal{R}_i|$. Let I be the index of the first ray \mathcal{R}_i for which $|\mathcal{R}_i| \leq \sqrt{2}\log^{1+\xi} n$, so that the last of this decreasing sequence of sectors is S_{I-1} .

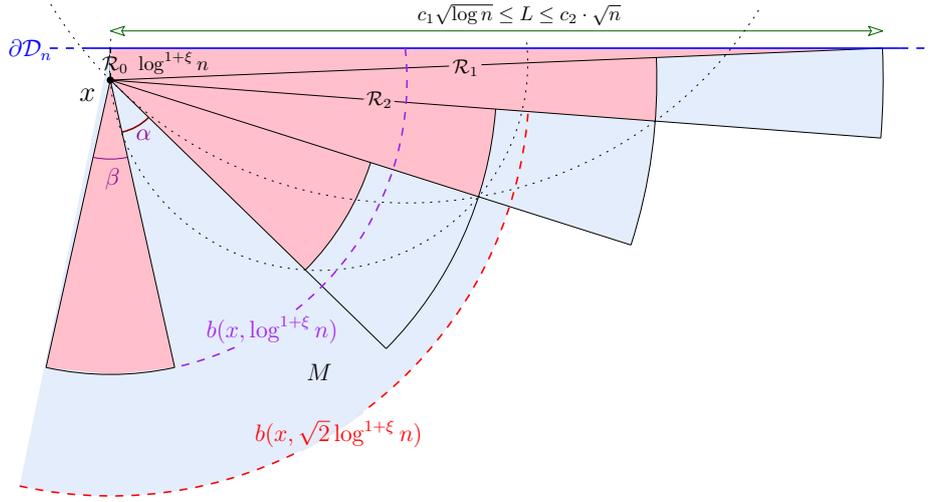


Figure 10: The construction to bound the maximum degree of a vertex near the border. The pink shaded circular sectors are each conditioned to contain at least one point. In this case, no Delaunay neighbour to the right of x can lie outside of the

Finally, we add a sector S_I directly opposite \mathcal{R}_0 within a ball of radius $\log^{1+\xi} n$. We choose its internal angle β so that its area is $\log^{1+\xi} n$.

We now proceed to showing that each circular sector enclosed by two adjacent rays contains a point with high probability and that when this event occurs, we have a bound on the region of points that may be a neighbour to x in $\text{DT}(\Phi)$.

Lemma 31. *Let α be the angle between the final ray \mathcal{R}_I and the edge of the sector S_{I+1} opposite \mathcal{R}_0 . Then, for n large enough, α is positive. This implies that the sectors S_i , $0 \leq i \leq I$ are disjoint.*

Proof. No angle between any two rays may exceed $2\beta = 4\pi \log^{-(1+\xi)} n$ since the area of a every sector is $\log^{1+\xi} n$, and the minimum circular radius of a sector is $\frac{1}{\sqrt{2}} \log^{1+\xi} n$. Also, the angle between the last ray \mathcal{R}_I and \mathcal{R}_1 is smaller than

$$\sum_{j=0}^{\infty} \frac{2\beta}{2^j} \leq 4\beta.$$

Since for n large enough, the angle $\mathcal{R}_0\mathcal{R}_1$ is smaller than $\frac{\pi}{2} + \frac{\pi}{12}$ we obtain $\alpha \geq \frac{\pi}{2} - \frac{\pi}{12} - \frac{\beta}{2} - 4\beta$ which is positive for n large enough. \square

Let $M \subset \mathbb{R}^2$ be defined by

$$M := b(x, \sqrt{2} \log^{1+\xi} n) \cup \bigcup_{i=1}^I Q_i,$$

that is the union of all shaded regions in Figure 10 (pink or blue).

Lemma 32. *Suppose that every for every $0 \leq i \leq I + 1$, we have $|\Phi \cap S_i| > 0$, and that the domain is rotated so that \mathcal{R}_0 is in exactly in the direction of the y -axis. Then every neighbour of x in $\text{DT}(\Phi)$ having positive x -coordinate of must lie in M .*

Proof. Any neighbour y of x in $\text{DT}(\Phi)$ must have a circle not containing any point of Φ that touches both x and y . Let $y \notin M$. Since y has positive x -coordinate, it is easy to see that any circle touching both x and y must also fully contain one of the sectors S_i , $0 \leq i \leq I$, in pink in Figure 10 (see dotted circles in Figure 10). By assumption, each S_i , $0 \leq i \leq I$ contains at least a site of Φ , so no circle touching both x and y can be empty, and y cannot be a neighbour of x in $\text{DT}(\Phi)$. \square

Lemma 33. *For any $\xi > 0$,*

$$\mathbb{P}\left(\max\left\{\delta_{\Phi}(x) \mid x \in \Phi, \|\partial\mathcal{D}_n x\| \leq \sqrt{\log n}\right\} \geq \log^{2+3\xi} n\right) \leq 2 \exp\left(-\log^{1+\xi} n\right),$$

for n large enough.

Proof. Let X_0 be chosen uniformly at random among the points of Φ within distance $\sqrt{\log n}$ of $\partial\mathcal{D}_n$. Note that if there is no such point, then

$$\max\left\{\delta_{\Phi}(x) : x \in \Phi, \|\partial\mathcal{D}_n x\| \leq \sqrt{\log n}\right\} = 0.$$

Let A_{X_0} be the event that no (pink) sector S_i , $0 \leq i \leq I$ about X_0 is empty and let B_{X_0} be the event that no sector S_i , $0 \leq i \leq I$ about X_0 contains more than $\log^{1+2\xi} n$ points (see Figure 10). Given that the number of sectors $I + 2$ about X_0 is deterministically bounded by $\frac{\pi}{6} \log^{1+\xi} n \leq n$ (for n large enough), we have by the union bound that the probability that

$$\mathbb{P}(A_{X_0}^c) \leq n \cdot \exp\left(-\log^{1+\xi} n\right) \quad \text{and} \quad \mathbb{P}(B_{X_0}^c) \leq n \cdot \exp\left(-\log^{1+2\xi} n\right).$$

Conditional on A_{X_0} and B_{X_0} occurring, we may count the number of points that could possibly be Delaunay neighbours of X_0 . This includes all points in the at most $\frac{\pi}{6} \log^{1+\xi} n$ sectors, each containing at most $\log^{1+2\xi} n$ points (conditional on B_{X_0}). We also add all the points not contained within any sector, but lying within the circle of radius $\sqrt{2} \log^{1+\xi}$ about X_0 (shaded blue in Figure 10). Standard arguments give that this region contains no more than $2\pi \log^{2+3\xi} n$ points with probability bounded at most $\exp\left(-\log^{2+3\xi} n\right)$. Putting these together and applying the union bound we have

$$\begin{aligned} & \mathbb{P}(\max\left\{\delta_{\Phi}(x) : x \in \Phi, \|\partial\mathcal{D}_n x\| \leq \sqrt{\log n}\right\} > \log^{2+3\xi} n) \\ & \leq 2n \cdot \mathbb{P}(\delta_{\Phi}(X_0) > \log^{2+3\xi} n) + \mathbb{P}(|\Phi| > 2n) \\ & \leq 2 \exp\left(-\log^{1+\xi} n\right), \end{aligned}$$

for n sufficiently large. \square

The second case of our proof is much simpler, since it suffices to bound the maximum distance between any two Delaunay neighbours, and then count the maximum number of points falling within this region.

Lemma 34. For any $\xi > 0$,

$$\mathbb{P}\left(\max\left\{\delta_{\Phi}(x) \mid x \in \Phi, \|\partial\mathcal{D}_n x\| > \sqrt{\log n}\right\} \geq \log^{2+\xi} n\right) \leq \exp\left(-\log^{1+\xi/3} n\right),$$

for n large enough.

Proof. Let $X_0 \in \Phi$ be chosen uniformly at random among the points of Φ at distance more than $\sqrt{\log n}$ from $\partial\mathcal{D}_n$. Again, note that if there is no such point,

$$\max\left\{\delta_{\Phi}(x) : x \in \Phi, \|\partial\mathcal{D}_n x\| > \sqrt{\log n}\right\} = 0.$$

If X_0 is well-defined, any neighbour of X_0 in $\text{DT}(\Phi)$ outside of the ball $b(X_0, 1/2 \log^{1/2+\xi} n)$ implies the existence of large region of \mathcal{D}_n that is empty of points of Φ . By adapting the proof of Lemma 19 and using Lemma 10, the probability that such a neighbour exists may be bounded by $\exp(-\log^{1+\xi} n)$. We omit the details.

We now upper bound the number of points that may fall within this ball. Split the domain into a regular grid with cells of side length $\frac{1}{2} \log^{1/2+\xi} n$. The probability that any of these grid cells contains more than $\log^{2+\xi/3}$ points is bounded by $\exp(-\log^{2+2\xi} n)$ for large n , and our ball may intersect at most four of these, so the degree of X_0 in $\text{DT}(\Phi)$ is bounded by $4 \log^{2+2\xi} n$ with probability $2 \exp(-\log^{1+\xi} n)$ in this case. The result follows from the union bound, just as in Lemma 33. \square

Acknowledgments We would like to thank Marc Glisse, Mordecai Golin, Jean-François Marckert, and Andrea Sportiello for fruitful discussions during the Presage workshop on geometry and probability.

References

- [1] F. Baccelli and B. Błaszczyszyn. *Stochastic Geometry and Wireless Networks*. NOW, 2009.
- [2] M. W. Bern, D. Eppstein, and F. F. Yao. The expected extremes in a delaunay triangulation. In *ICALP*, pages 674–685, 1991.
- [3] N. Bonichon and J.-F. Marckert. Asymptotics of geometrical navigation on a random set of points in the plane. *Adv. in Applied Probability*, 43:899–942, 2011.
- [4] C. Bordenave. Navigation on a Poisson point process. *The Annals of Applied Probability*, 18:708–746, 2008.
- [5] P. Bose and L. Devroye. On the stabbing number of a random Delaunay triangulation. *Computational Geometry: Theory and Applications*, 36:89–105, 2006.
- [6] P. Bose and P. Morin. Online routing in triangulations. *SIAM journal on computing*, 33: 937–951, 2004.

- [7] P. Bose and P. Morin. Competitive online routing in geometric graphs. *Theor. Comput. Sci.*, 324(2-3):273–288, 2004.
- [8] P. Bose, A. Brodnik, S. Carlsson, E. Demaine, R. Fleischer, A. López-Ortiz, P. Morin, and J. Munro. Online routing in convex subdivisions. *International Journal of Computational Geometry & Applications*, 12:283–295, 2002.
- [9] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities - A nonasymptotic theory of independence*. Clarendon Press, 2012.
- [10] D. Chen, L. Devroye, V. Dujmović, and P. Morin. Memoryless routing in convex subdivisions: Random walks are optimal. *Computational Geometry*, 45(4):178 – 185, 2012.
- [11] P. M. M. de Castro and O. Devillers. Simple and efficient distribution-sensitive point location, in triangulations. In *ALLENEX*, pages 127–138, 2011.
- [12] O. Devillers, S. Pion, and M. Teillaud. Walking in a triangulation. *Int. J. Found. Comput. Sci.*, 13:181–199, 2002.
- [13] L. Devroye, E. Mücke, and B. Zhu. A note on point location in Delaunay triangulations of random points. *Algorithmica*, 22:477–482, 1998.
- [14] L. Devroye, C. Lemaire, and J.-M. Moreau. Expected time analysis for Delaunay point location. *Comput. Geom. Theory Appl.*, 29:61–89, 2004.
- [15] D. Dubhashi and A. Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009.
- [16] S. Janson. Large deviation for sums of partially dependent random variables. *Random Structures and Algorithms*, 24(3):234–248, 2004.
- [17] G. Kozma, Z. Lotker, and M. Sharir. Geometrically aware communication in random wireless networks. *ACM PODC*, 2004.
- [18] C. L. Lawson. Software for C^1 surface interpolation. In J. R. Rice, editor, *Math. Software III*, pages 161–194. Academic Press, New York, NY, 1977.
- [19] X.-Y. Li, G. Calinescu, P.-J. Wan, and Y. Wang. Localized delaunay triangulation with application in ad hoc wireless networks. *IEEE Trans. Parallel Distrib. Syst.*, 14(10):1035–1047, 2003.
- [20] J. Matoušek. *Lectures on discrete geometry*. Springer, 2002.
- [21] C. J. H. McDiarmid. Concentration. In Habib et al., editor, *Probabilistic Methods in Algorithmic Discrete Mathematics*, pages 195–248. Springer-Verlag, 1998.
- [22] S. Misra, I. Woungang, and S. Misra. *Guide to wireless sensor networks*. Springer, 2009.

- [23] E. P. Mücke, I. Saias, and B. Zhu. Fast randomized point location without preprocessing in two- and three-dimensional delaunay triangulations. *IN PROC. 12TH ANNU. ACM SYMPOS. COMPUT. GEOM*, 12:274–283, 1996.
- [24] Y. Wang and X.-Y. Li. Efficient delaunay-based localized routing for wireless sensor networks: Research articles. *Int. J. Commun. Syst.*, 20(7):767–789, 2007.
- [25] G. Xia. Improved upper bound on the stretch factor of Delaunay triangulations. In *Proceedings of the 27th annual ACM symposium on Computational geometry (SoCG)*, pages 264–273. ACM, 2011.
- [26] G. Xia and L. Zhang. Toward the tight bound of the stretch factor of delaunay triangulations. In *CCCG*, 2011.
- [27] F. Zhao and L. Guibas. *Wireless Sensor Networks. An Information Processing Approach*. Morgan Kaufmann, 2004.
- [28] B. Zhu. On Lawson’s oriented walk in random Delaunay triangulations. In *Fundamentals of Computation Theory*, volume 2751 of *Lecture Notes Comput. Sci.*, pages 222–233. Springer-Verlag, 2003.