From Proximity to Utility: A Voronoi Partition of Pareto Optima*

Hsien-Chih Chang† Sariel Har-Peled‡ Benjamin Raichel§

August 8, 2014

Abstract

We present an extension of Voronoi diagrams where not only the distance to the site is taken into account when considering which site the client is going to use, but additional attributes (i.e., prices or weights) are also considered. A cell in this diagram is then the loci of all clients that consider the same set of sites to be relevant. In particular, the precise site a client might use from this candidate set depends on parameters that might change between usages, and the candidate set lists all of the relevant sites. The resulting diagram is significantly more expressive than Voronoi diagrams, but naturally has the drawback that its complexity, even in the plane, might be quite high.

Nevertheless, we show that if the attributes of the sites are drawn from the same distribution (note that the locations are fixed), then the expected complexity of the candidate diagram is near linear. To derive this result, we derive several new technical results, which are of independent interest.

1. Introduction

Informal description of the candidate diagram. Suppose you open your refrigerator one day, to discover it is time to go grocery shopping.[1] Which store you go to will be determined by a number of different factors. For example, what items you are buying, and do you want the cheapest price or highest quality, and how much time you have for this chore. Naturally the distance to the store will also be a factor. On different days which store is the best to go to will differ based on that day's preferences. However, there are certain stores you will never shop at. These are stores which are worse in every way than some other store (i.e., further, more expensive, lower quality, etc.). Therefore, the stores that are relevant and therefore in the *candidate set* are those that are not strictly worse in every way than some other store. Thus, every point in the plane is mapped

*Work on this paper was partially supported by NSF AF award CCF-1421231, and CCF-1217462. The paper is also available on the arXiv [CHR14].

†Department of Computer Science, University of Illinois; 201 N. Goodwin Avenue, Urbana, IL 61801, USA; hchang17@illinois.edu; web.engr.illinois.edu/~hchang17.

‡Department of Computer Science, University of Illinois; 201 N. Goodwin Avenue, Urbana, IL 61801, USA; sariel@uiuc.edu; sarielhp.org.

§Department of Computer Science, University of Illinois; 201 N. Goodwin Avenue, Urbana, IL 61801, USA; raichel2@uiuc.edu; illinois.edu/~raichel2.

[1]Unless you are feeling adventurous enough that day to eat the frozen mystery food stuck to the back of the freezer, which we strongly discourage you from doing.

to a set of stores that a client at that location might use. The *candidate diagram* is the partition of the plane into regions, where each candidate set is the same for all points in the same region. Naturally, if your only consideration is distance, then this is the (classical) Voronoi diagram of the sites. However, here, deciding which shop to use is an instance of multi-objective optimization — as there are multiple, potentially competing, objectives to be optimized, and the decision might change as the weighting and influence of these objectives mutate over time (in particular, you might decide to do your shopping in several of these stores for different products). The concept of relevant stores discussed above is often refereed as the *Pareto optima*.

Pareto optima in welfare economics. Pareto efficiency, named after Vilfredo Pareto, is a core concept in economic theory and more specifically in welfare economics. Here each point in $\mathbb{R}^d$ represents the corresponding utilities of d players for a particular allocation of finite resources. A point is said to be *Pareto optimal* if there is no other allocation which increases the utility of any individual without decreasing the utility of another. The *First Fundamental Theorem of Welfare Economics* states that any competitive equilibrium (i.e., supply equals demand) is Pareto optimal. The origins of this theorem date back to 1776 with Adam Smith's famous (and controversial) work, "The Wealth of Nations," but was not formally *proven* until the 20th century by Lerner, Lange, and Arrow (see [Fel08]). Naturally such proofs rely on simplifying (i.e., potentially unrealistic) assumptions such as perfect knowledge, or absence of externalities. The *Second Fundamental Theorem of Welfare Economics* states that any Pareto optimum is achievable through lump-sum transfers (i.e. taxation and redistribution). In other words each Pareto optima is a "best solution" under some set of societal preferences, and is achievable through redistribution in one form or another (see [Fel08] for a more in depth discussion).

Pareto optima in computer science. In computational geometry such Pareto optima points relate to the *orthogonal convex hull* [OSW84], which in turns relates to the well known convex hull (the input points that lie on the orthogonal convex hull is a super set of those which lie on the convex hull). Pareto optima are also of importance to the database community [BKS01, HTC13], in which context such points are called *maximal* or *skyline points*. Such points are of interest as they can be seen as the relevant subset of the (potentially much larger) result of a relational database query. The standard example is that of querying a database of hotels for the cheapest and closest hotel, where naturally hotels which are farther and more expensive than an alternative hotel are not relevant results. There is a significant amount of work on computing these points, see Kung *et al.* [KLP75]. More recently, Godfrey *et al.* [GSG07] compared various approaches for the computation of these points (from a databases perspective), as well as introduce their own new external algorithm.[2]

Modeling uncertainty. Recently, there is a growing interest in modeling uncertainty in data. As real data is acquired via physical measurements, noise and errors are introduced. This can be addressed by treating the data as coming from a distribution (e.g., a point location might be interpreted as a center of a Gaussian), and computing desired classical quantities adapted for such settings. Thus, a nearest-neighbor query becomes a probabilistic question — what is the expected

[2]There is of course a lot of other work on Pareto optimal points, from connections to Nash equilibrium to scheduling. We resisted the temptation of including many such references which are not directly related to our paper.

distance to the nearest-neighbor? What is the most likely point to be the nearest-neighbor? (See [AAH+13] and references therein for more information.)

This in turn gives rise to the question of what is the expected complexity of geometric structures defined over such data. The case where the data is a set of points, and the locations of the points are chosen randomly was thoroughly investigated (see [SW93, WW93, HR14] and references therein). The problem when the locations are fixed but weights associated with the points are chosen randomly is relatively newer. Agarwal *et al.* [AHKS13] showed that for a set of disjoint segments in the plane, if they are being expanded randomly, then the expected complexity of the union is near linear. This result is somewhat surprising as in the worst case the complexity of such a union is quadratic.

Our result can be interpreted as bounding the complexity of the level set of multiplicatively weighted Voronoi diagrams, where the weights are being chosen randomly. Here, the weight of a site can model, for example, the price of delivery of a single unit as a linear function of the delivery distance (in reality, the pricing function might be significantly more involved, and our model, described below, captures such situations). Har-Peled and Raichel [HR14] gave a simpler proof that works for the whole diagram, that with randomly weighted points in the plane the complexity of the diagram is $O(n \operatorname{polylog} n)$ in expectation. The current write-up is a continuation of this work, to handle more general kinds of weighting function.

1.1. Our contributions

Conceptual contribution. We define formally the *candidate diagram* in Section 2.2 — a new geometric structure that combines proximity information with utility. For every point x in the plane, the diagram associates a *candidate set* $\mathsf{L}(\mathsf{x})$ of sites that are relevant to x; that is, all the sites that are Pareto optima for x. Putting it differently, a site is irrelevant to x (that is, not in $\mathsf{L}(\mathsf{x})$) if it is further away and worse in all parameters than some other site. Significantly, unlike the traditional Voronoi diagram, the candidate diagram allows the user to change their distance function, as long as the function respects the domination relationship.

This diagram is a significant extension of the Voronoi diagram, and includes other extensions of Voronoi diagrams as special subcases, like multiplicative weighted Voronoi diagrams. Not surprisingly, the worst case complexity of this diagram can be quite high.

Technical contribution. We consider the case where each site chooses its jth attribute from some distribution $\mathcal{D}_j$ independently for each j. We show that the candidate diagram in expectation has near linear complexity, and that, with high probability, the candidate set has polylog size for any point in the plane. In the process we derive several results which are interesting in their own right.

(A) **Low complexity of the minima for random points in the hypercube.** We prove that if n points are sampled from a fixed distribution (see Section 2.3 for assumptions on the distribution) over the d-dimensional hypercube then, with probability $1 - 1/n^{\Omega(1)}$, the number of Pareto optima points is $O(\log^{d-1} n)$, which is optimal up to a constant factor with the expectation (see Lemma 5.4).
Previously, this result was only known in a weaker form that is insufficient to imply our other results. Specifically, Bai *et al.* [BDHT05] proved that after normalization the cumulative distribution function of number of Pareto optima points is normal, up to an additive error $O(1/\operatorname{polylog} n)$. (See [BR10b, BR10a] as well.) In particular, their results (which

are quite nice and mathematically involved) can only imply the statement with probability $1 - 1/\text{polylog}\, n$.

To the best of our knowledge this result is new — we emphasize, however, that for our purposes a weaker bound of $O(\log^d n)$ is sufficient, and such a weaker result follows readily from the ε-net theorem (naturally, this would add a log factor to later results in the paper).

(B) **Backward analysis with high probability.** To get this result, we prove a lemma providing high probability bounds when applying backwards analysis [Sei93] (see Lemma 7.1). Such tail estimates are known in the context of randomized incremental algorithms [CMS93, BCKO08], but our proof is arguably more direct and cleaner, and should be applicable to more cases. See Section 2.4 and Section 7.

(C) **Overlay of the kth order Voronoi cells in randomized incremental construction.** We prove that the overlay of cells during a randomized incremental construction of the kth order Voronoi diagram is of complexity $O(k^4 n \log n)$ (see Lemma 4.8).

(D) **Complexity of the candidate diagram.** Combining the above results carefully, and inserting the sites in the right order, yields a near-linear upper bound on the complexity of the candidate diagram (see Theorem 6.1).

Context. In addition to the conceptual and technical contributions listed above, our new results suggest that while the candidate diagram might have prohibitive complexity in the worse case, for real world inputs their complexity might be significantly smaller. In particular, the expressiveness of the candidate diagram potentially makes it a useful data structure for real world applications.

1.2. Outline of the paper

In Section 2 after some required preliminaries (Section 2.1), we formally introduce the candidate diagram (Section 2.2), describe our sampling model (Section 2.3), and discuss backward analysis (Section 2.4). To bound the complexity of the candidate diagram (i.e., both the planar partition, and the set of all distinct candidate sets), in Section 3 we define and analyze, an enlarged candidate set called the *proxy set* (Section 3.1); we bound the size of the proxy set using backward analysis (Section 3.2), then prove that with high probability, the proxy set contains the candidate set (Section 3.3). In Section 4 we show that the appropriate diagram for this enlarged set is the overlay of cells during the randomized incremental construction of the kth order Voronoi diagram (Section 4.1); using the Clarkson-Shor moment technique (Section 4.2), we can provide an expected bound of $O(k^4 n \log n)$ on the complexity of the diagram (Section 4.3). Next, in Section 5, by analyzing the number of staircase points of random points in hypercubes (Section 5.1), we bound the expected complexity of the candidate set for any point in the plane (Section 5.2). In Section 6 we prove our main result, showing the desired bound on the complexity of the candidate diagram. In Section 7, we complete our proof by providing high-probability bound for backward-analysis.

2. The candidate diagram

2.1. Preliminaries

Throughout, we assume the reader is familiar with standard computational geometry terms, such as arrangements [SA95], vertical-decomposition [BCKO08], etc. In the same vein, we assume that variable d, the *dimension*, is a small constant and the O notation hides constants that are potentially exponential (or worse) in d.

A quantity is bounded by $O(f)$ ***with high probability***, denoted as $O_{\mathrm{whp}}(f)$, if for any large enough constant γ, there is another constant c depending on γ such that the quantity is at most $c \cdot f$ with probability at least $1 - n^{-\gamma}$. In other words, the bound holds for any sufficiently small polynomial error with the expense of a multiplicative constant factor on the size of the bound.

Definition 2.1. Consider two points $\mathsf{p} = (\mathsf{p}_1, \ldots, \mathsf{p}_d)$ and $\mathsf{q} = (\mathsf{q}_1, \ldots, \mathsf{q}_d)$ in $\mathbb{R}^d$. The point p ***dominates*** q (denoted by $\mathsf{p} \preceq \mathsf{q}$) if $\mathsf{p}_i \leq \mathsf{q}_i$, for all i.

Given a point set $\mathsf{P} \subseteq \mathbb{R}^d$, there are several terms for the subset of P that is not dominated, as already discussed above, such as *Pareto optima* or *minima*. Here, we use the following term.

Definition 2.2. For a point set $\mathsf{P} \subseteq \mathbb{R}^d$, a point $\mathsf{p} \in \mathsf{P}$ is a ***staircase point*** of P if no other point of P dominates it. The set of all such points, denoted by $\mathsf{St}(\mathsf{P})$, is the ***staircase*** of P.

2.2. Formal definition of the candidate diagram

Let $\mathsf{S} = \{\mathsf{s}_1, \ldots, \mathsf{s}_n\}$ be a set of n ***sites*** in the plane. For each site s in S, there is an associated list $\alpha = \langle \alpha_1, \ldots, \alpha_d \rangle$ of d real-valued attributes, each in the interval $[0, 1]$. When viewed as a point in the unit hypercube $[0, 1]^d$, this list of attributes is the ***parametric point*** of the site s_i. Specifically, a site is a point in the plane encoding a facility location, while the term *point* is used to refer to the (parametric) point encoding its attributes in $\mathbb{R}^d$.

Preferences. Fix a client location x in the plane. For each site, there are $d+1$ associated variables for the client to consider. Specifically, the client distance to the site, and d additional attributes (e.g., prices of d different products) associated with the site. Conceptually, the goal of the client is to "pay" as little as possible by choosing the best site (e.g., minimize the overall cost of buying these d products together from a site, where the price of traveling the distance to the site is also taken into account).

Definition 2.3. A client x has a ***dominating preference*** if for any two sites s, s' in the plane, with parametric points α, $\alpha' \in \mathbb{R}^d$ respectively, the client would prefer the site s over s' if $\|\mathsf{x} - \mathsf{s}\| \leq \|\mathsf{x} - \mathsf{s}'\|$ and $\alpha \preceq \alpha'$ (that is, α dominates α').

Note that a client having a dominating preference does not identify a specific optimum site for the client, but rather a set of potential optimum sites. Specifically, given a client location x in the plane, let its distance to the ith site be $\ell_i = \|\mathsf{x} - \mathsf{s}_i\|$. The set of sites the client might possibly use (assuming the client uses a dominating preference) are the staircase points of the set $\mathsf{P}(\mathsf{x}) = \{(\alpha_1, \ell_1), \ldots, (\alpha_n, \ell_n)\}$ (i.e., we are adding the distance to each site as an additional attribute of the site — this attribute depends on the location of x). The set of sites realizing the staircase of $\mathsf{P}(\mathsf{x})$ (i.e., all the sites relevant to x) is the ***candidate set*** $\mathsf{L}(\mathsf{x})$ of x:

$$\mathsf{L}(\mathsf{x}) = \left\{ \mathsf{s}_i \in \mathsf{S} \;\middle|\; (\alpha_i, \ell_i) \text{ is a staircase point of } \mathsf{P}(\mathsf{x}) \text{ in } \mathbb{R}^{d+1} \right\}. \tag{2.1}$$

The ***candidate cell*** of x is the set of all the points in the plane that have the same candidate set associated with them; that is, $\{\mathsf{p} \in \mathbb{R}^2 \mid \mathsf{L}(\mathsf{p}) = \mathsf{L}(\mathsf{x})\}$. The decomposition of the plane into these cells is the ***candidate diagram***.

Now, the client x has the candidate set $\mathsf{L}(\mathsf{x})$, and it chooses some site (or potentially several sites) from $\mathsf{L}(\mathsf{x})$ that it might want to use. Note that the client might decide to use different sites for different acquisitions.

Complexity of the diagram. The ***complexity*** of a planar arrangement is the total number of edges, faces, and vertices. A candidate diagram can be interpreted as a planar arrangement, and its complexity is defined analogously. The ***space complexity*** of the candidate diagram is the total amount of memory needed to store the diagram explicitly, and is bounded by the complexity of the candidate diagram together with the sum of the sizes of candidate sets over all the faces in the arrangement of the diagram (which is potentially larger by a factor of n, the number of sites). Note, that the space complexity is a somewhat naïve upper bound, as using persistent data-structures might significantly reduce the space needed to store the candidate lists.

Lemma 2.4. *Given n sites in the plane, the complexity of the candidate diagram is $O(n^4)$. The space complexity of the candidate diagram is $\Omega(n^2)$ and $O(n^5)$.*

Proof: The lower bound is easy, and is left as an exercise to those interested reader. A naïve upper bound of $O(n^5)$ on the space complexity, follows from the fact that (i) all possible pairs of sites induce together $\binom{n}{2}$ bisectors, (ii) the complexity of the arrangement of the bisectors is $O(n^4)$, and (iii) the candidate set of each face in this arrangement might have n elements inside. ■

We leave the question of closing the gap in the bounds of Lemma 2.4 as an open problem for further research.

2.3. Sampling model

Fortunately, the situation changes when randomization is involved. Let S be a set of n sites in the plane. For each site $\mathsf{s} \in \mathsf{S}$, a parametric point $\alpha = (\alpha_1, \ldots, \alpha_d)$ is sampled independently from $[0, 1]^d$, with the following constrain: each coordinate α_i is sampled from a (continuous) distribution $\mathcal{D}_i$ for each i independently as well. In particular, the sorted order of the n parametric points by a specific coordinate yields a uniform random permutation (for the sake of simplicity of exposition we assume that all the values sampled are distinct).

Our main result shows that, under the above assumptions, both the complexity and the space complexity of the candidate diagram is near-linear in expectation — see Theorem 6.1 for the exact statement.

2.4. A short detour into backward analysis

Randomized incremental construction is a powerful technique used by geometric algorithms. Here, one is given a set of elements S (e.g., segments in the plane), and one is interested in computing some structure induced by these elements (e.g., the vertical decomposition formed by the segments). To this end, one computes a random permutation $\Pi = \langle \mathsf{s}_1, \ldots, \mathsf{s}_n \rangle$ of the elements of S, and in the ith iteration one computes the structure V_i induced by the ith prefix $\Pi_i = \langle \mathsf{s}_1, \ldots, \mathsf{s}_i \rangle$ of Π, by inserting the ith element s_i into V_{i-1} and updating into V_i (e.g., split all the vertical trapezoids of V_{i-1} that intersect s_i, and merge together adjacent trapezoids with the same floor and ceiling).

In ***backward analysis*** one is interested in computing the probability that a specific object that exists in V_i was actually created in the ith iteration (e.g., a specific vertical trapezoid in the vertical decomposition V_i). Observing that assuming the object of interest is defined by at most b elements of Π_i for some constant b implies that the desired probability is equal to the probability that s_i is one of these defining elements, which is at most b/i. In some cases, one is interested in the summation of these probabilities over the n iterations (which is $O(b \log n)$), as way of counting the number of times certain events happen during the incremental construction. However, this yields only a bound in expectation. For a high probability bound, one can not apply this argument directly, as there is a subtle dependency leakage between the corresponding indicator variables involved between different iterations. (Without going into a detailed example, this is because the defining sets of the objects of interest can have different sizes, and these sizes depend on which elements were used in the permutation in earlier iterations.)

Let P be a set of n elements. A ***property*** $\mathcal{P}$ is a function that maps any subset X of P to a subset $\mathcal{P}(\mathsf{X})$ of X. Intuitively the elements in $\mathcal{P}(\mathsf{X})$ have some desired property with respect to X (for example, let X be a set of points in the plane, then $\mathcal{P}(\mathsf{X})$ may be those points who lie on the convex hull of X). The following corollary to Lemma 7.1 provides a high probability bound for backward analysis, and while the proof is an easy application of the Chernoff inequality, it nevertheless significantly simplifies some classical results on randomized incremental construction algorithms. See Section 7 for a more detailed discussion and a proof.

Corollary 2.5. *Let P be a set of n elements, and $k \geq 1$ be a fixed integer depending only on P, and let $\mathcal{P}(\mathsf{X})$ be a property defined over any subset $\mathsf{X} \subseteq \mathsf{P}$. Now, consider a random permutation $\langle \mathsf{p}_1, \ldots, \mathsf{p}_n \rangle$ of P, and let $\mathsf{P}_i = \{\mathsf{p}_1, \ldots, \mathsf{p}_i\}$ be the underlying set of the prefix $\langle \mathsf{p}_1, \ldots, \mathsf{p}_i \rangle$. Furthermore, assume that for each i we have $|\mathcal{P}(\mathsf{P}_i)| \leq k$. Let X_i be the indicator variable of the event $\mathsf{p}_i \in \mathcal{P}(\mathsf{P}_i)$. Then, for any constant $\gamma \geq 2e$, we have*

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i > \gamma \cdot (2k \ln n)\right] < n^{-\gamma k}.$$

3. The proxy set

Providing a reasonable bound on the complexity of the candidate diagram directly seems challenging. Therefore, we instead define for each point x in the plane a slightly different set, called the *proxy set*. First we prove that the size of the proxy set for every point in the plane has small size (see Lemma 3.2 below); then we prove that, with high probability, the proxy set of x contains the candidate set of x for all points in the plane (see Lemma 3.4 below).

3.1. Definitions

As before, the input is a set of sites S. For each site $\mathsf{s} \in \mathsf{S}$, we randomly pick a parametric point $\alpha \in [0,1]^d$ according to the sampling method described in Section 2.3.

Volume ordering. Given a point $\mathsf{p} = (\mathsf{p}_1, \ldots, \mathsf{p}_d)$ in $[0,1]^d$, the ***point volume*** $\mathrm{pv}(\mathsf{p})$ of point p is defined to be $\mathsf{p}_1 \mathsf{p}_2 \cdots \mathsf{p}_d$; that is, the volume of the hyperrectangle with p and the origin as a pair of opposite corners. When p is specifically the associated parametric point of an input site s, we refer to the point volume of p as the ***parametric volume*** of s. Observe that if point p dominates another point q then p must have smaller point volume (i.e., p lies in the hyperrectangle defined by q).

The ***volume ordering*** of sites in S is a permutation $\langle \mathsf{s}_1, \ldots, \mathsf{s}_n \rangle$ ordered by increasing parametric volume of the sites; that is, $\mathrm{pv}(\alpha_1) \leq \mathrm{pv}(\alpha_2) \leq \ldots \leq \mathrm{pv}(\alpha_n)$, where α_i is the parametric point of s_i. If α_i dominates α_j then s_i precedes s_j in the volume ordering. So if we add the sites in volume ordering, then when we add the ith site s_i we can ignore all later sites when determining its region of influence (i.e., the region of points whose candidate set s_i belongs to), as no later points can dominate s_i.

k nearest neighbors. For a set of sites S and a point x in the plane, let $\mathsf{d}_k(\mathsf{x}, \mathsf{S})$ denote the ***kth nearest neighbor distance*** to x in S; that is, the kth smallest value in the multiset $\left\{ \|\mathsf{x} - \mathsf{s}\| \;\middle|\; \mathsf{s} \in \mathsf{S} \right\}$. The ***k nearest neighbors*** to x in S is the set

$$\mathsf{N}_k(\mathsf{x}, \mathsf{S}) = \left\{ \mathsf{s} \in \mathsf{S} \;\middle|\; \|\mathsf{x} - \mathsf{s}\| \leq \mathsf{d}_k(\mathsf{x}, \mathsf{S}) \right\}.$$

Definition 3.1. Let S be a set of sites in the plane, and let $\mathsf{V} = \langle \mathsf{s}_1, \ldots, \mathsf{s}_n \rangle$ be the volume ordering of S. Denote S_i as the underlying set of the ith prefix $\langle \mathsf{s}_1, \ldots, \mathsf{s}_i \rangle$ of V. For a parameter k and a point x in the plane, the ***kth proxy set*** of x is the set of sites

$$\mathsf{C}_k(\mathsf{x}, \mathsf{V}) = \bigcup_{i=1}^{n} \mathsf{N}_k(\mathsf{x}, \mathsf{S}_i),$$

In words, site s is in $\mathsf{C}_k(\mathsf{x}, \mathsf{V})$ if it is one of the k nearest neighbors to point x in some prefix of V.

3.2. Bounding the size of the proxy set

The desired bound now follows by using backward analysis and Corollary 2.5.

Lemma 3.2. *Let S be a set of n sites in the plane and V be the volume ordering of S, and let $k \geq 1$ be a fixed parameter. Then we have $|\mathsf{C}_k(\mathsf{x}, \mathsf{V})| = O_{whp}(k \log n)$ simultaneously for all points x in the plane.*

Proof: Fix a point x in the plane. Ordering the sites by increasing parametric volume creates a random permutation V on the distances of the sites from x. A site s gets added to the kth proxy set $\mathsf{C}_k(\mathsf{x}, \mathsf{V})$ if site s is one of the k nearest neighbors of x among the underlying set S_i of some prefix of V. Therefore a direct application of Corollary 2.5 implies (by setting $\mathcal{P}(\mathsf{S}_i)$ to be $\mathsf{N}_k(\mathsf{x}, \mathsf{S}_i)$), with high probability, that $\left|\mathsf{C}_k(\mathsf{x}, \mathsf{V})\right| = O(k \log n)$.

Furthermore, this holds for all points in the plane simultaneously. Indeed, consider the arrangement determined by the $\binom{n}{2}$ bisectors formed by all the pairs of sites in S. This arrangement is a simple planar map with $O(n^4)$ vertices and $O(n^4)$ faces. Observe that within each face the proxy set cannot change since all points in this face have the same ordering of their distances to the sites in S. Therefore, picking a representative point from each of these $O(n^4)$ faces, applying the high probability bound to each one of them, and then applying the union bound implies the claim. ■

3.3. The proxy set contains the candidate set

The following corollary is implied by a careful (but straightforward) integration argument (see Appendix A).

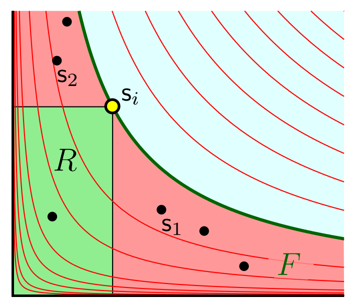
Corollary 3.3. *Let $F_d(\Delta)$ be the total measure of the points $\mathsf{p} \in [0,1]^d$, such that the point volume $\mathrm{pv}(\mathsf{p})$ is at most Δ. Then for $\Delta \geq (\log n)/n$ we have $F_d(\Delta) = \Theta(\Delta \log^{d-1} n)$; in particular, $F_d(\log n/n) = \Theta((\log^d n)/n)$.*

Lemma 3.4. *Let S be a set of n sites in the plane and V be the volume ordering of S, and let $k = \Theta(\log^d n)$ be a fixed parameter. For all points x in the plane, we have that $\mathsf{L}(\mathsf{x}) \subseteq \mathsf{C}_k(\mathsf{x}, \mathsf{V})$, and this holds with high probability.*

Proof: Fix a point x in the plane, and let s_i be any site *not* in $\mathsf{C}_k(\mathsf{x}, \mathsf{V})$, and let α_i be the associated parametric point. We claim that, with high probability, the site s_i is dominated by some other site which is closer to x, and hence by the definition of dominating preference (Definition 2.3), s_i cannot be a site used by x (and thus $\mathsf{s}_i \notin \mathsf{L}(\mathsf{x})$). Taking the union bound over all sites not in $\mathsf{C}_k(\mathsf{x}, \mathsf{V})$ then implies this claim.

By Corollary 3.3, the total measure of the points in $[0,1]^d$ with point volume at most $\Delta = \log n/n$ is $\Theta((\log^d n)/n)$. As such, by Chernoff's inequality, with high probability, there are $K = O(\log^d n)$ sites in S that their parametric points have point volume smaller than Δ. In particular, by choosing k to be sufficiently large (i.e., $k > K$), the underlying set S_k of the kth prefix of V will contain all these small point volume sites, and since $\mathsf{S}_k \subseteq \mathsf{C}_k(\mathsf{x}, \mathsf{V})$, so will $\mathsf{C}_k(\mathsf{x}, \mathsf{V})$. Therefore, from this point on, we will assume that $\mathsf{s}_i \notin \mathsf{C}_k(\mathsf{x}, \mathsf{V})$ and $\Delta_i = \mathrm{pv}(\alpha_i) = \Omega(\log n/n)$.

Now any site s with smaller parametric volume than s_i are all in the (unordered) prefix S_i. In particular, the k nearest neighbors $\mathsf{N}_k(\mathsf{x}, \mathsf{S}_i)$ of x in S_i all have smaller parametric volume than s_i. Hence $\mathsf{C}_k(\mathsf{x}, \mathsf{V})$ contains k points all of which have smaller parametric volume than s_i, and which are closer to x. Therefore, the claim will be implied if one of these k points dominates s_i.



Let $\Delta_i = \mathrm{pv}(\alpha_i)$. Now, the probability of a site s (that is closer to x than s_i) with parametric point α to dominate s_i is the probability that $\alpha \preceq \alpha_i$ given that $\alpha \in F$, where $F = \{\alpha \in [0,1]^d \mid \mathrm{pv}(\alpha) \leq \Delta_i\}$. By Corollary 3.3, we have $\mathrm{vol}(F) = F_d(\Delta_i) = \Theta(\Delta_i(\log^{d-1} n))$. The probability that a random parametric point picked in $[0,1]^d$ dominates α_i is exactly $\mathrm{pv}(\alpha_i)$ (i.e., $\mathrm{area}(R)$), and as such the desired probability $\mathbf{Pr}\left[\alpha \preceq \alpha_i \mid \alpha \in F\right] = \Delta_i / F_d(\Delta_i) = O(1/\log^{d-1} n)$. This is depicted in the figure on the right — the probability of a random point picked uniformly from the region F under the curve $y = \Delta_i/x$, induced by s_i, to fall in the rectangle R.

As the parametric point of each one of the k points in $\mathsf{N}_k(\mathsf{x}, \mathsf{S}_i)$ has equal probability to be anywhere in F, this implies the expected number of points in $\mathsf{N}_k(\mathsf{x}, \mathsf{S}_i)$ which dominate s_i is $\mathbf{Pr}\left[\alpha \preceq \alpha_i \mid \alpha \in F\right] \cdot k = \Theta(\log n)$. Therefore by making k sufficiently large, Chernoff's inequality implies the desired result.

It follows that this holds, for all points in the plane simultaneously, by following the argument used in the proof of Lemma 3.2. ∎

4. Bounding the complexity of the kth order proxy diagram

The ***kth proxy cell*** of x is the set of all the points in the plane that have the same kth proxy set associated with them; that is, $\{\mathsf{p} \in \mathbb{R}^2 \mid \mathsf{C}_k(\mathsf{p}, \mathsf{V}) = \mathsf{C}_k(\mathsf{x}, \mathsf{V})\}$. The decomposition of the plane into these faces is the ***kth order proxy diagram***. In this section, our goal is to prove that the expected total diagram complexity of the kth order proxy diagram is $O\bigl(k^4 n \log n\bigr)$.

To this end, we bound the complexity of the kth order proxy diagram by relating it to the overlay of star-polygons that rise out of the kth order Voronoi diagram.

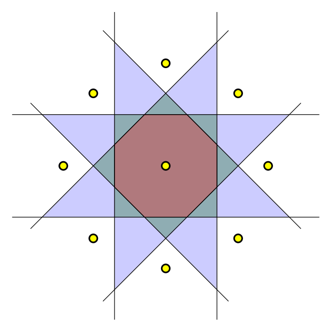
4.1. Preliminaries

4.1.1. The kth order Voronoi diagram

Let S be a set of n sites in the plane. The ***kth order Voronoi diagram*** of S is a partition of the plane into faces such that each cell is the locus of points which have the same set of k nearest sites of S (the internal ordering of these k sites, by distance to the query point, may vary within the cell however). It is well known that the worst case complexity of this diagram is $\Theta(k(n-k))$ (see [AKL13, Section 6.5]).

Environments and overlays. For a site s in S and a constant k, the ***k environment*** of s, denoted by $\mathrm{env}_k(\mathsf{s}, \mathsf{S})$, is the set of all the points in the plane such that s is one of their k nearest neighbors in S; that is,

$$\mathrm{env}_k(\mathsf{s}, \mathsf{S}) = \left\{ \mathsf{x} \in \mathbb{R}^2 \;\middle|\; \mathsf{s} \in \mathsf{N}_k(\mathsf{x}, \mathsf{S}) \right\}.$$



See the figure on the right for an example how this environment looks like for different values of k. One can view the k environment of s as the union of the kth order Voronoi cells which have s as one of the k nearest sites. Observe that the overlay of the polygons $\mathrm{env}_k(\mathsf{s}_1, \mathsf{S}), \ldots, \mathrm{env}_k(\mathsf{s}_n, \mathsf{S})$ produces the kth order Voronoi diagram of S. Indeed, for any point x in the plane, if s is one of x's k nearest sites, then by definition x is covered by $\mathrm{env}_k(\mathsf{s}, \mathsf{S})$; and conversely if x is covered by $\mathrm{env}_k(\mathsf{s}, \mathsf{S})$ then s is one of x's k nearest neighbors. It is also known that each k environment of a site is a star-shaped polygon; this was previously observed by Aurenhammer and Schwarzkoff [AS92]. For the sake of completeness we include a proof here.

Lemma 4.1. *The set* $\mathrm{env}_k(\mathsf{s}, \mathsf{S})$ *is a star shaped polygon with respect to the point* s.

Proof: Consider the set of all $n-1$ bisectors determined by s and any other site in S. For any point x in the plane, $\mathsf{p} \in \mathrm{env}_k(\mathsf{s}, \mathsf{S})$ holds if the segment from s to p crosses at most $k-1$ of these bisectors. The star-shaped property follows as when walking along any ray emanating from s, the number of bisectors crossed is a monotonically increasing function of distance from s. Moreover, $\mathrm{env}_k(\mathsf{s}, \mathsf{S})$ is a polygon as its boundary is composed of subsets of straight line bisectors. ■

Going back to our original problem, let k be a fixed constant, and let $\mathsf{V} = \langle \mathsf{s}_1, \ldots, \mathsf{s}_n \rangle$ be the volume ordering of S. As usual, we use S_i to denote the unordered ith prefix of V. Let $\mathrm{env}_i := \mathrm{env}_k(\mathsf{s}_i, \mathsf{S}_i)$, that is, the union of all the cells in the kth order Voronoi diagram of S_i where s_i is one of the k nearest neighbors.

Observation. *The arrangement determined by the overlay of the polygons* $\mathrm{env}_1, \ldots, \mathrm{env}_n$ *is the* k*th order proxy diagram of* S.

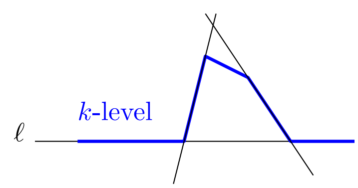
4.1.2. Arrangements of planes and lines

One can interpret the kth order Voronoi diagram in terms of an arrangement of planes in $\mathbb{R}^3$. Specifically, lift each site to the paraboloid defined by the set of points $(x, y, x^2 + y^2)$. Consider the arrangement of planes tangent to the paraboloid at the lifted locations of the sites. A point on the union of these planes is of *level* k if there are exactly k planes strictly below it. The ***k-level*** is the closure of the set of points of level k. Let $\mathsf{E}_k(\mathsf{S})$ denote the set of edges in the arrangement on the k-level, where an edge is a maximal portion of the k-level that lies on the intersection of two planes (induced by two sites). Consider a point x in the xy-plane. The increasing z-ordering of the planes vertically above x is the same as the ordering, by increasing distance from x, to the corresponding sites. Hence, the projection of the edges in $\mathsf{E}_{k-1}(\mathsf{S})$ onto the xy-plane results in the edges of the kth order Voronoi diagram. The set of all the edges from the first k levels is denoted by $\mathsf{E}_{\leq k}(\mathsf{S})$.

For a set of lines L in the plane, one can define the k-levels similarly. We again use $\mathsf{E}_k(L)$ to denote the set of edges in the arrangement of L on the k-level. We need the following lemmas.

Lemma 4.3. *Let L be a set of lines in general position in the plane, and let ℓ be any line in L. Then at most $k + 2$ edges from $\mathsf{E}_k(L)$, the k-level of the arrangement of L, can lie on ℓ.*

Proof: This lemma is well known, and its proof is included here for the sake of completeness.



Perform a linear transformation such that ℓ is horizontal and the k-level is preserved. As we go from left to right along the now horizontal line ℓ (starting at infinity), we may leave and enter the k-level multiple times. However, every time we leave and then return to the k-level we must intersect a negative slope line. Specifically, both when we leave and return to the k-level, there must be an intersection with another line. If when leaving, this intersection is with a negative slope line then we are done, so assume it has positive slope. In this case the level on ℓ decreases as we leave the k-level, therefore when we return to the k-level, the point of return must be at an intersection with a negative slope line (since only negative slope intersections can increase the level), see figure on the right.

So after leaving and returning to the k-level $k + 1$ times, there must be at least $k + 1$ negative slope lines below, which implies that the remaining part of ℓ is on level strictly larger than k. ■

Lemma 4.4. *Let L be a set of n lines in general position in the plane. Fix any arbitrary insertion ordering of the lines in L, and let m be the total number of distinct vertices on the k-level of the arrangement of L seen over all iterations of this insertion process. We have $m = O(nk)$.*

Proof: Let ℓ_i be the ith line inserted, and let L_i be the set of the first i inserted lines. Any new vertex on the kth level created by the insertion must lie on ℓ_i. However, by Lemma 4.3 at most $k + 2$ edges from $\mathsf{E}_k(L_i)$ can lie on ℓ_i. As each such edge has at most two endpoints, the insertion of ℓ_i contributes $O(k)$ vertices to the k-level. The bound now follows by summing over all n lines. ■

4.2. Bounding the size of the below conflict-lists

4.2.1. The below conflict lists

Let H be a set of n planes in general position in $\mathbb{R}^3$. (For example, in the setting of kth order Voronoi diagram, H is the set of planes that are tangent to the paraboloid at the lifted locations of the sites.) For any subset $\mathsf{R} \subseteq \mathsf{H}$, let $\mathsf{V}_k(\mathsf{R})$ denote the vertices on the k-level of the arrangement of R. Similarly, let $\mathsf{V}_{\leq k}(\mathsf{R}) = \bigcup_{i=0}^{k} \mathsf{V}_k(\mathsf{R})$ be the set of vertices of level at most k in the arrangement of R, and let $\mathsf{E}_{\leq k}(\mathsf{R})$ be the set of edges of level at most k in the arrangement of R. For a vertex v in the arrangement of R, the ***below conflict list*** of v contains those planes in H that lie strictly below v; denote b_v to be $|\mathsf{B}(v)|$. For an e in the arrangement of R, let $\mathsf{B}(e)$ be the set of planes of H which lie below e (i.e., there is at least one point on e that lies above such a plane); denote b_e to be $|\mathsf{B}(e)|$. Our purpose here is to bound the quantity $\mathbf{E}\left[\sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} \mathsf{b}_v\right]$.

4.2.2. The Clarkson-Shor technique

In the following, we use the Clarkson-Shor technique [CS89], stated here without proof (see [Har11] for details). Specifically, let S be a set of elements such that any subset $\mathsf{R} \subseteq \mathsf{S}$ defines a corresponding set of objects $\mathcal{T}(\mathsf{R})$ (e.g., S is a set of planes and any subset $\mathsf{R} \subseteq \mathsf{S}$ induces the set of vertices in the arrangement of planes). Each potential object, τ, has a defining set and a stopping set. The ***defining set***, $D(\tau)$, is a subset of S that must appear in R in order for the object to be present in $\mathcal{T}(\mathsf{R})$. We require that the defining set has at most a constant size for every object. The ***stopping set***, $\kappa(\tau)$, is a subset of S such that if any of its member appear in R then τ is not present in $\mathcal{T}(\mathsf{R})$. We also naturally require that $\kappa(\tau) \cap D(\tau) = \emptyset$ for all object τ. Surprisingly, this already implies the following.

Theorem 4.5 (Bounded Moments, [CS89]). *Using the above notation, let S be a set of n elements, and let R be a random sample of size r from S. Let $f(\cdot)$ be a polynomially growing function[3]. We have that*

$$\mathbf{E}\left[\sum_{\tau \in \mathcal{T}(\mathsf{R})} f\Big(|\kappa(\tau)|\Big)\right] = O\Big(\mathbf{E}\Big[|\mathcal{T}(\mathsf{R})|\Big]\, f\Big(\frac{n}{r}\Big)\Big),$$

where the expectation is taken over random sample R.

4.2.3. Bounding the below conflict-lists

The technical challenge. The proof of the next lemma is technically interesting as it does not follow in a straightforward fashion from the Clarkson-Shor technique. Indeed, the below conflict list is *not* the standard conflict list. Specifically, the decision whether a vertex v in the arrangement of R is of level at most k is a "global" decision of R, and as such the defining set of this vertex is neither of constant size, nor unique, as required to use the Clarkson-Shor technique. If this was the only issue, the extension by Agarwal *et al.* [AMS98] could handle this situation. However it is even worse: a plane $h \in \mathsf{H} \setminus \mathsf{R}$ that is below a vertex $v \in \mathsf{V}_{\leq k}(\mathsf{R})$ is not necessarily conflicting with v (i.e., in the stopping set of v) — as its addition to R will not necessarily remove v from $\mathsf{V}_{\leq k}(\mathsf{R} \cup \{h\})$.

The solution. Since the standard techniques fails in this case, we need to perform our argument somehow indirectly. Specifically, we use a second random sample and then deploy the Clarkson-Shor technique on this smaller sample — this is reminiscent of the proof bounding the size of $\mathsf{V}_{\leq k}(\mathsf{H})$

[3] A function $f(n)$ is a ***polynomially growing*** function, if (i) $f(\cdot)$ is monotonically increasing, (ii) for any integers $i, n \geq 1$, $f(i \cdot n) = i^{O(1)} f(n)$. This holds for example if $f(n)$ is a constant degree polynomial of n, with all its coefficients being positive. Of course, it holds for a much larger family of functions, e.g. $f(i) = i \log i$.

by Clarkson-Shor [CS89], and the proof of the exponential decay lemma of Chazelle and Friedman [CF90].

Lemma 4.6. *Let k be a fixed constant, and let R be a random sample (without replacement) of size r from a set of H of n planes in $\mathbb{R}^3$, we have* $\mathbf{E}\Big[\sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} \mathsf{b}_v\Big] = O\big(nk^3\big)$.

Proof: From the sake of simplicity of exposition, let us assume that the sampling here is done by picking every element into the random sample R with probability r/n. Doing the computations below using sampling without replacement (so we get the exact size) requires modifying the calculations so that the probabilities are stated using binomial coefficients — this makes the calculation messier, but the results remain the same. See [Sha03] for further discussion of this minor issue.

So, fix a sample R and sample each plane in R, with probability $1/k$, to be in R'. Let us consider the probability that a vertex $v \in \mathsf{V}_{\leq k}(\mathsf{R})$ ends up on the lower envelope of R'. A lower bound can be achieved by the standard argument of Clarkson-Shor. Specifically, if a vertex v is on the lower envelope then its three defining planes must be in R' and moreover as $v \in \mathsf{V}_{\leq k}(\mathsf{R})$ by definition there are at most k planes below it that must not be in R'. So let X_v be an indicator variable for whether v appears on the lower envelope of R', we then have

$$\mathbf{E}_{\mathsf{R}'}\big[X_v \bigm| \mathsf{R}\big] \geq \frac{1}{k^3}(1 - 1/k)^k \geq \frac{1}{e^2 k^3}.$$

Observe that

$$\mathbf{E}_{\mathsf{R}'}\Big[\sum_{v \in \mathsf{V}_0(\mathsf{R}')} \mathsf{b}_v\Big] = \mathbf{E}_{\mathsf{R}}\Big[\mathbf{E}_{\mathsf{R}'}\Big[\sum_{v \in \mathsf{V}_0(\mathsf{R}')} \mathsf{b}_v \Bigm| \mathsf{R}\Big]\Big] = \mathbf{E}_{\mathsf{R}}\Big[\mathbf{E}_{\mathsf{R}'}\Big[\sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} X_v \mathsf{b}_v \Bigm| \mathsf{R}\Big]\Big]. \tag{4.1}$$

Fixing the value of R, the lower bound above implies

$$\mathbf{E}_{\mathsf{R}'}\Big[\sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} X_v \mathsf{b}_v \Bigm| \mathsf{R}\Big] = \sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} \mathbf{E}_{\mathsf{R}'}\Big[X_v \mathsf{b}_v \Bigm| \mathsf{R}\Big] = \sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} \mathsf{b}_v \mathbf{E}_{\mathsf{R}'}\Big[X_v \Bigm| \mathsf{R}\Big] \geq \sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} \frac{\mathsf{b}_v}{e^2 k^3},$$

by linearity of expectations and as b_v is a constant for v. Plugging this into Eq. (4.1), we have

$$\mu = \mathbf{E}_{\mathsf{R}'}\left[\sum_{v \in \mathsf{V}_0(\mathsf{R}')} \mathsf{b}_v\right] \geq \mathbf{E}_{\mathsf{R}}\left[\sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} \frac{\mathsf{b}_v}{e^2 k^3}\right] = \frac{1}{e^2 k^3} \mathbf{E}_{\mathsf{R}}\left[\sum_{v \in \mathsf{V}_{\leq k}(\mathsf{R})} \mathsf{b}_v\right]. \tag{4.2}$$

Observe that R' is a random sample of R which by itself is a random sample of H. As such, one can interpret R' as a direct random sample of H. The lower envelope of a set of planes has linear complexity, and for a vertex v on the lower envelope of R' the set $\mathsf{B}(v)$ is the standard conflict list of v. As such, Theorem 4.5 implies

$$\mu = \mathbf{E}_{\mathsf{R}'}\Big[\sum_{v \in \mathsf{V}_0(\mathsf{R}')} \mathsf{b}_v\Big] = O\left(|\mathsf{R}'| \cdot \frac{n}{|\mathsf{R}'|}\right) = O(n).$$

Plugging this into Eq. (4.2) implies the claim. ∎

Corollary 4.7. *Let R be a random sample (without replacement) of size r from a set H of n planes in $\mathbb{R}^3$. We have that* $\mathbf{E}_{\mathsf{R}}\Big[\sum_{e \in \mathsf{E}_{\leq k}(\mathsf{R})} \mathsf{b}_e\Big] = O\big(nk^3\big)$.

Proof: Under general position assumption every vertex in the arrangement of H is adjacent to 8 edges. For an edge $e = uv$, it is easy to verify that $\mathsf{B}(e) \subseteq \mathsf{B}(u) \cup \mathsf{B}(v)$, and as such we charge the conflict list of e to its two endpoints u and v, and every vertex get charged $O(1)$ times. Now, the claim follows by Lemma 4.6.

This argument fails to capture edges that are rays in the arrangement, but this is easy to overcome by clipping the arrangement to a bounding box that contains all the vertices of the arrangement. We omit the easy but tedious details. ■

4.3. Putting it all together

The proof of the following lemma is similar in spirit to the argument in [HR14].

Lemma 4.8. *Let S be a set of n sites in the plane, $\mathsf{V} = \langle \mathsf{s}_1, \ldots, \mathsf{s}_n \rangle$ be a random permutation of S, and let k be a fixed number. The expected complexity of arrangement determined by the overlay of the polygons $\mathrm{env}_1, \ldots, \mathrm{env}_n$ (and therefore, the expected complexity of the kth order proxy diagram) is $O\big(k^4 n \log n\big)$, where $\mathrm{env}_i = \mathrm{env}_k(\mathsf{s}_i, \mathsf{S}_i)$ and $\mathsf{S}_i = \{\mathsf{s}_1, \ldots, \mathsf{s}_i\}$ is the underlying set of the ith prefix of V, for each i.*

Proof: As the arrangement of the overlay of the polygons $\mathrm{env}_1, \ldots, \mathrm{env}_n$ is a planar map it suffices to bound the number of edges in the arrangement. For each i, let $\mathsf{E}(\mathrm{env}_i)$ be the edges in $\mathsf{E}_{\leq k}(\mathsf{S}_i)$ that appear on the boundary of env_i (for simplicity we do not distinguish between edges in $\mathsf{E}_{\leq k}(\mathsf{S}_i)$ in $\mathbb{R}^3$ and their projection in the plane). Created in the ith iteration, an edge e in $\mathsf{E}(\mathrm{env}_i)$ is going to be broken into several pieces in the final arrangement of the overlay. Let n_e be the number of such pieces that arise from e.

Fix an integer i. As S_i is fixed, $\mathsf{B}(e)$ is also fixed, for all $e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)$. Moreover, we claim that $n_e \leq c \cdot k \mathsf{b}_e$ for some constant c. Indeed, n_e counts the number of future intersections of e with the edges of $\mathsf{E}(\mathrm{env}_j)$, for any $j > i$. As the edge e is on the k-level at the time of creation, and the edges in $\mathsf{E}(\mathrm{env}_j)$ are on the k-level when they are being created (in the future), these edges must lie below e. Namely, any future intersect on e are caused by intersections of (pairs of) planes in $\mathsf{B}(e)$. So consider the intersection of all planes in $\mathsf{B}(e)$ on the vertical plane containing e. On this vertical plane, $\mathsf{B}(e)$ is a set of b_e lines, whose insertion ordering is defined by the suffix of the permutation $\langle \mathsf{s}_{i+1}, \ldots, \mathsf{s}_n \rangle$. Now any edge of $\mathsf{E}(\mathrm{env}_j)$, for some $j > i$, that intersects e must appear as a vertex on the k-level at some point during the insertion of these lines. However, by Lemma 4.4, applied to the lines of $\mathsf{B}(e)$ on the vertical plane of e, under any insertion ordering there are at most $O(k \mathsf{b}_e)$ vertices that ever appear on the k-level.

For an edge $e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)$, let X_e be the indicator variable of the event that e was created in the ith iteration, and furthermore, lies on the boundary of env_i. Observe that $\mathbf{E}[X_e \mid \mathsf{S}_i] \leq 4/i$, as an edge appears for the first time in round i only if one of its (at most) four defining sites was the ith site inserted.

Let $Y_i = \sum_{e \in \mathsf{E}(\mathrm{env}_i)} n_e = \sum_{e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)} n_e X_e$ be the total (forward) complexity contribution to the final arrangement of edges added in round i. We thus have

$$\mathbf{E}\Big[Y_i \;\Big|\; \mathsf{S}_i\Big] = \mathbf{E}\Bigg[\sum_{e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)} n_e X_e \;\Bigg|\; \mathsf{S}_i\Bigg] \leq \mathbf{E}\Bigg[\sum_{e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)} ck \mathsf{b}_e X_e \;\Bigg|\; \mathsf{S}_i\Bigg] = \sum_{e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)} ck \mathsf{b}_e \, \mathbf{E}\Big[X_e \;\Big|\; \mathsf{S}_i\Big]$$
$$\leq \frac{4ck}{i} \sum_{e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)} \mathsf{b}_e.$$

The total complexity of the overlay arrangement of the polygons $\mathrm{env}_1, \ldots, \mathrm{env}_n$ is asymptotically bounded by $\sum_i Y_i$, and so by Corollary 4.7 we have

$$\mathbf{E}\Bigl[\sum_i Y_i\Bigr] = \sum_i \mathbf{E}\Bigl[\mathbf{E}\Bigl[Y_i \Bigm| \mathsf{S}_i\Bigr]\Bigr] \leq \sum_i \mathbf{E}\left[\frac{4ck}{i} \sum_{e \in \mathsf{E}_{\leq k}(\mathsf{S}_i)} \mathsf{b}_e\right] = O\left(\sum_i \frac{nk^4}{i}\right) = O\Bigl(k^4 n \log n\Bigr). \qquad \blacksquare$$

5. On the expected size of the staircase

5.1. Number of staircase points

5.1.1. The two dimensional case

Let P be a set of n points sampled uniformly at random from the unit square. If we order the points in P by increasing x-coordinate, then the staircase points are exactly the points which have the smallest y-values out of all points in their prefix in this ordering. As the x-coordinates are sampled uniformly at random, this ordering is a random permutation $\langle y_1, \ldots, y_n \rangle$ of the y-values Y. Let X_i be the indicator variable of the event that y_i is the smallest number in $\mathsf{Y}_i = \{y_1, \ldots, y_i\}$ for each i. By setting property $\mathcal{P}(\mathsf{Y}_i)$ to be the smallest number in the prefix Y_i, we have $\sum_{i=1}^n X_i = O(\log n)$ with high probability by Corollary 2.5.

Corollary 5.1. *Let P be a set of n points sampled uniformly at random from the unit square $[0,1]^2$. Then the number of staircase points $\mathsf{St}(\mathsf{P})$ in P is $O_{whp}(\log n)$.*

5.1.2. Higher dimensions

Lemma 5.2. *Let m and n be parameters, such that $m \leq n$. Let $\mathsf{Q} = \langle \mathsf{q}_1, \ldots, \mathsf{q}_m \rangle$ be an ordered set of m points picked randomly from $[0,1]^d$ as described in Section 2.3. Assume that we have $|\mathsf{St}(\mathsf{Q}_i)| = O_{whp}(c_d \log^{d-1} n)$ for every i, where $\mathsf{Q}_i = \{\mathsf{q}_1, \ldots, \mathsf{q}_i\}$ is underlying set of the ith prefix of Q. Then, the set $\mathsf{S} = \bigcup_{i=1}^m \mathsf{St}(\mathsf{Q}_i)$ has size $O_{whp}(c_d \log^d n)$.*

Proof: By setting $\mathcal{P}(\mathsf{Q}_i) = \mathsf{St}(\mathsf{Q}_i)$, by Corollary 2.5 we have that $\mathbf{Pr}\bigl[|\mathsf{S}| > \gamma(2k \ln m)\bigr] \leq m^{-\gamma k}$ with $k = O(c_d \log^{d-1} n)$, for any $\gamma \geq 2e$. Setting $\gamma = O((\log n)/(\log m))$ then implies the claim. $\blacksquare$

Lemma 5.3. *Fix a dimension $d \geq 2$. Let m, n be parameters, such that $m \leq n$. Let P be a set of m points picked randomly from $[0,1]^d$ as described in Section 2.3. Then, $|\mathsf{St}(\mathsf{P})| = O_{whp}(c_d \log^{d-1} n)$ holds for some constant c_d that depends only on d.*

Proof: The argument follows by induction on dimension. The two-dimensional case follows from Corollary 5.1. Assume we have proven the claim for all dimension smaller than d.

Now, sort P by increasing value of the dth coordinate, and let $\mathsf{p}_i = (\mathsf{q}_i, \ell_i)$ be the ith point in P in this order for each i, where q_i is a $(d-1)$-dimensional vector and ℓ_i is the value of the dth coordinate of p_i. Observe that the points $\mathsf{q}_1, \ldots, \mathsf{q}_m$ are randomly, uniformly, and independently picked from the hypercube $[0,1]^{d-1}$. Now, if p_i is a minima point of P, then it is a minima point of $\{\mathsf{p}_1, \ldots, \mathsf{p}_i\}$. But this implies that q_i is a minima point of $\mathsf{Q}_i = \{\mathsf{q}_1, \ldots, \mathsf{q}_i\}$ as well; namely, $\mathsf{q}_i \in \mathsf{S} := \bigcup_{i=1}^m \mathsf{St}(\mathsf{Q}_i)$. This implies that $|\mathsf{St}(\mathsf{P})| \leq |\mathsf{S}|$. Now, applying induction hypothesis on each Q_i in dimension $d-1$ we have $|\mathsf{St}(\mathsf{Q}_i)| = O\bigl(c_{d-1} \log^{d-2} n\bigr)$ holds for all i, with high probability, and plugging it into Lemma 5.2 we have $|\mathsf{St}(\mathsf{P})| \leq |\mathsf{S}| = O\bigl(c_{d-1} \log^{d-1} n\bigr)$, with high probability. Choosing a proper constant c_d now implies the claim. $\blacksquare$

Lemma 5.4. *Fixed a dimension $d \geq 2$. Let $\mathsf{Q} = \langle \mathsf{q}_1, \ldots, \mathsf{q}_n \rangle$ be an ordered set of n points picked randomly from $[0,1]^d$ (as described in Section 2.3), and $\mathsf{Q}_i = \{\mathsf{q}_1, \ldots, \mathsf{q}_i\}$ is the ith (unordered) prefix of Q. Then, the set $\bigcup_{i=1}^{n} \mathsf{St}(\mathsf{Q}_i)$ is of size $O_{whp}(c_d \log^d n)$, and the staircase $\mathsf{St}(\mathsf{P})$ is of size $O_{whp}(c_d \log^{d-1} n)$.*

Proof: By Lemma 5.2, the set $\bigcup_{i=1}^{n} \mathsf{St}(\mathsf{Q}_i)$ is of size $O\big(c_d \log^d n\big)$, with high probability. By Lemma 5.3, the set $\mathsf{St}(\mathsf{P})$ is of size $O\big(c_d \log^{d-1} n\big)$, with high probability. ∎

Remark. In the proof of Lemma 5.3 whether a point is on the staircase (or not) only depends on the coordinate orderings of the points and not their actual values.

The basic recursive argument used in Lemma 5.3 was used by Clarkson [Cla04] to bound the expected number of k-sets for a random point set. Here, using Corollary 2.5 enables us to get a high-probability bound.

Note that the definition of the staircase can be made with respect to any corner of the hypercube (i.e., this corner would replace the origin in the definition dominance, point volume, the exponential grid, etc.). Taking the union over all 2^d such staircases gives us the subset of P on the orthogonal convex hull of P. Therefore Lemma 5.4 also bounds the number of input points on the orthogonal convex hull. As the vertices on the convex hull of P are a subset of the points in P on the orthogonal convex hull, the above also implies the same bound on the number of vertices on the convex hull.

5.2. Bounding the size of the candidate set

We can now readily bound the size of the candidate set for any point in the plane.

Lemma 5.6. *Let S be a set of n sites in the plane, where for each site s in S a parametric point from a distribution over $[0,1]^d$ as described in Section 2.3. Then, the candidate set has size $O_{whp}(\log^d n)$ simultaneously for all points in the plane.*

Proof: Consider the arrangement of bisectors of all pairs of points of S. This arrangement has complexity $O\big(n^4\big)$; inside each cell the candidate set is the same. Now for any point in a cell of the arrangement, Lemma 5.4 immediately gives us the stated bound, with high probability. Therefore picking a representative point from each cell in this arrangement and applying the union bound imply the claim. ∎

6. The main result

We now use the bound on the complexity of the proxy diagram, as well as our knowledge of the relationship between the candidate set and the proxy set to bound the complexity of the candidate diagram.

Theorem 6.1. *Let S be a set of n sites in the plane, where for each site in S we sample an associated parametric point in $[0,1]^d$ in the way described in Section 2.3. Then the expected complexity of the candidate diagram is $O\big(n \log^{8d+5} n\big)$; the expected space complexity of the candidate diagram is $O\big(n \log^{9d+5} n\big)$.*

Proof: Let $\mathsf{V} = \langle \mathsf{s}_1, \ldots, \mathsf{s}_n \rangle$ be the volume ordering of S, and fix the value of k to be sufficiently large, such that $k = \Theta\big(\log^d n\big)$. By Lemma 4.8 the expected complexity of the proxy diagram is

$O\big(k^4 n \log n\big)$. Triangulating each polygonal cell in the diagram does not increase its asymptotic complexity. Lemma 3.2 implies that, (simultaneously) for all the points in the plane, the proxy set has size $O(k \log n)$, with high probability. Now, Lemma 3.4 implies that, with high probability, for any point in the plane, the proxy set contains the candidate set.

The resulting triangulation has $O\big(k^4 n \log n\big)$ faces, and inside each face all the sites that might appear in the candidate set are all present in the proxy set of this face. By Lemma 2.4, the complexity of an m-site candidate diagram is $O\big(m^4\big)$. Therefore the complexity of the candidate diagram per face is $O\big((k \log n)^4\big)$ (clipping the candidate diagram of these sites to the containing triangle does not increase the asymptotic complexity). Multiplying by the number of faces, $O\big(k^4 n \log n\big)$, by the complexity of the arrangement within each face, $O\big((k \log n)^4\big)$, yields the desired result.

The bound on the space complexity follows readily from the bound on the size of the candidate set from Lemma 5.6. ■

7. Backward analysis with high probability

Lemma 7.1. *Let P be a set of n elements, and $k \geq 1$ be a fixed integer depending only on P. Let $\mathcal{P}(\mathsf{X})$ be a property defined over any subset $\mathsf{X} \subseteq \mathsf{P}$ (see Section 2.4), satisfying the following condition: if $|\mathsf{X}| \geq k$ then $|\mathcal{P}(\mathsf{X})| = k$. Now, consider a random permutation $\langle \mathsf{p}_1, \ldots, \mathsf{p}_n \rangle$ of P. For each i, denote $\mathsf{P}_i = \{\mathsf{p}_1, \ldots, \mathsf{p}_i\}$ and let X_i be an indicator variable of the event $\mathsf{p}_i \in \mathcal{P}(\mathsf{P}_i)$. Then we have*

- *(A)* $\mathbf{Pr}\Big[\sum_{i=1}^{n} X_i > \gamma \cdot (2k \ln n)\Big] \leq n^{-\gamma k}$ *for any $\gamma \geq 2e$.*
- *(B) The bound in (A) holds under a weaker condition: for all $\mathsf{X} \subseteq \mathsf{P}$ we have that $|\mathcal{P}(\mathsf{X})| \leq k$.*
- *(C) The bound in (A) holds under an even weaker condition: For each i, we have $|\mathcal{P}(\mathsf{P}_i)| \leq k$ (i.e., we only need bounds for those prefix sets).*

Proof: (A) Let E_i be the event that $\mathsf{p}_i \in \mathcal{P}(\mathsf{P}_i)$. It suffices to show that the events $E_1, \ldots, E_n$ are mutually independent. The insight is to think about the sampling process of creating the random permutation of P in a different way, and then the result readily follows. Indeed, we randomly pick a permutation of the given elements, and set the last element to be p_n. Next, pick a random permutation of the remaining elements and set the last element as the second-to-last element (i.e., p_{n-1}) in the output permutation. Repeat this process until the whole permutation is generated. Observe that E_j is determined before E_i for any $j > i$.

Now, consider $1 \leq i_1 < i_2 < \ldots < i_t \leq n$, and observe that

$$\mathbf{Pr}\Big[E_{i_t} \;\Big|\; E_{i_1} \cap \ldots \cap E_{i_{t-1}}\Big] = \mathbf{Pr}\Big[E_{i_t}\Big]$$

since by our thought experiment, E_{i_t} is determined before all the other events $E_{i_{t-1}}, \ldots, E_{i_1}$. As such, we have by induction that

$$\begin{aligned}\mathbf{Pr}\Big[E_{i_1} \cap \ldots \cap E_{i_t}\Big] &= \mathbf{Pr}\Big[E_{i_t} \;\Big|\; E_{i_1} \cap \ldots \cap E_{i_{t-1}}\Big]\,\mathbf{Pr}\big[E_{i_1} \cap \ldots \cap E_{i_{t-1}}\big] \\ &= \mathbf{Pr}[E_{i_t}]\,\mathbf{Pr}\Big[E_{i_1} \cap \ldots \cap E_{i_{t-1}}\Big] = \prod_{j=1}^{t} \mathbf{Pr}\Big[E_{i_j}\Big],\end{aligned}$$

which implies that these events are mutually independent. Now

$$\mathbf{Pr}\Big[E_i\Big] = \min\!\left(\frac{|\mathcal{P}(\mathsf{P}_i)|}{i}, 1\right) = \min\!\left(\frac{k}{i}, 1\right),$$

and we have

$$\mu = \mathbf{E}\left[\sum_i X_i\right] \leq k + \sum_{i=k+1}^{n} \frac{k}{i} \leq k\left(1 + \ln n + 1 - \ln k\right) \leq k\left(2 + \ln \frac{n}{k}\right) \leq 2k \ln n.$$

For any constant $\delta \geq 2e$, by Chernoff's inequality, we have $\mathbf{Pr}\left[\sum_i X_i > \delta\mu\right] < 2^{-\delta\mu}$. Therefore by setting $\delta = \gamma(2k \ln n)/\mu$ (which is at least $2e$), we have

$$\mathbf{Pr}\left[\sum_i X_i > \gamma(2k \ln n)\right] = \mathbf{Pr}\left[\sum_i X_i > \frac{\gamma(2k \ln n)}{\mu}\mu\right] < 2^{-\gamma(2k \ln n)} < n^{-\gamma k}.$$

(B) In order the extend the result using the weaker condition, we augment the given property $\mathcal{P}$ to a new property $\mathcal{P}'$ that holds for *exactly* k elements. So, fix an arbitrary ordering on the elements of P. Now given any set X with $|\mathsf{X}| \geq k$, if $|\mathcal{P}(\mathsf{X})| = k$ then let $\mathcal{P}'(\mathsf{X}) = \mathcal{P}(\mathsf{X})$; otherwise, let $\mathcal{Q}$ be the $k - |\mathcal{P}(\mathsf{X})|$ smallest elements in $\mathsf{X} \setminus \mathcal{P}(\mathsf{X})$ according to the ordering on P, and let $\mathcal{P}'(\mathsf{X}) = \mathcal{P}(\mathsf{X}) \cup \mathcal{Q}$. The new property $\mathcal{P}'$ complies with the original condition. For any X, $\mathcal{P}(\mathsf{X}) \subseteq \mathcal{P}'(\mathsf{X})$, which implies that an upper bound on the probability that the ith element is in the property set $\mathcal{P}'$ is an upper bound on the corresponding probability for $\mathcal{P}$.

(C) Follows readily by observing that the required conditions on $\mathcal{P}$ applies only to the prefix sets $\mathsf{P}_1, \ldots, \mathsf{P}_n$. ■

The result of Lemma 7.1 is known in the context of randomized incremental construction algorithms (see [BCKO08, Section 6.4]). However, the known proof is more convoluted — indeed, if the property $\mathcal{P}(\mathsf{X})$ has different sizes for different sets X, then it is no longer true that variables X_i in the proof of Lemma 7.1 are independent. Thus the padding idea in part (B) of the proof is crucial in making the result more widely applicable.

Example. To see the power of Lemma 7.1 we provide two easy applications — both results are of course known, and are included here to make it clearer in what settings Lemma 7.1 can be applied. The impatient reader is encouraged to skip this example.

(A) **QuickSort**: We conceptually can think about **QuickSort** as being a randomized incremental algorithm, building up a list of numbers in the order they are used as pivots. Consider the execution of **QuickSort** when sorting a set P of n numbers. Let $\langle \mathsf{p}_1, \ldots, \mathsf{p}_n \rangle$ be the random permutation of the numbers picked in sequence by **QuickSort**. Specifically, in the ith iteration, it randomly picks a number p_i that was not handled yet, pivots based on this number, and then recursively handles the subproblems. At the ith iteration, a set $\mathsf{P}_i = \{\mathsf{p}_1, \ldots, \mathsf{p}_i\}$ of pivots has already been chosen by the algorithm. Consider a specific element $\mathsf{x} \in \mathsf{P}$. For any subset $\mathsf{X} \subseteq \mathsf{P}$, let $\mathcal{P}(\mathsf{X})$ be the two numbers in X having x in between them in the original ordering of P and are closest to each other. In other words, $\mathcal{P}(\mathsf{X})$ contains the (at most) two elements that are the endpoints of the interval of $\mathbb{R} \setminus \mathsf{X}$ that contains x. Let X_i be the indicator variable of the event $\mathsf{p}_i \in \mathcal{P}(\mathsf{P}_i)$ — that is, x got compared to the ith pivot when it was inserted. Clearly, the total number of comparisons x participates in is $\sum_i X_i$, and by Lemma 7.1 the number of such comparisons is $O(\log n)$, with high probability, implying that **QuickSort** takes $O(n \log n)$ time, with high probability.

(B) **Point-location queries in a history dag**: Consider a set of lines in the plane, and build their vertical decomposition using randomized incremental construction. Let $L_n = \langle \ell_1, \ldots, \ell_n \rangle$ be the permutation used by the randomized incremental construction. Given a query point p, the point-location time is the number of times the vertical trapezoid containing p changes in the vertical decomposition of $L_i = \langle \ell_1, \ldots, \ell_i \rangle$, as i increases. Thus, let X_i the indicator variable of the event that ℓ_i is one of the (at most) four lines defining the vertical trapezoid containing p the vertical decomposition of L_i. Again, Lemma 7.1 applies and implies that the query time is $O(\log n)$, with high probability. This result is well known, see [CMS93] and [BCKO08, Section 6.4], but our proof is arguably more direct and cleaner.

Acknowledgments

The authors would like to thank Pankaj Agarwal, Ken Clarkson, Nirman Kumar, and Raimund Seidel for useful discussions related to this work.

References

[AAH+13] P. K. Agarwal, B. Aronov, S. Har-Peled, J. M. Phillips, K. Yi, and W. Zhang. Nearest neighbor searching under uncertainty II. In *Proc. 32nd ACM Sympos. Principles Database Syst.* (PODS), pages 115–126, 2013.

[AHKS13] P. K. Agarwal, S. Har-Peled, H. Kaplan, and M. Sharir. Union of Random Minkowski Sums and Network Vulnerability Analysis. *ArXiv e-prints*, October 2013.

[AKL13] F. Aurenhammer, R. Klein, and D.-T. Lee. *Voronoi Diagrams and Delaunay Triangulations.* World Scientific, 2013.

[AMS98] P. K. Agarwal, J. Matoušek, and O. Schwarzkopf. Computing many faces in arrangements of lines and segments. *SIAM J. Comput.*, 27(2):491–505, 1998.

[AS92] F. Aurenhammer and O. Schwarzkopf. A simple on-line randomized incremental algorithm for computing higher order Voronoi diagrams. *Internat. J. Comput. Geom. Appl.*, pages 363–381, 1992.

[BCKO08] M. de Berg, O. Cheong, M. van Kreveld, and M. H. Overmars. *Computational Geometry: Algorithms and Applications.* Springer-Verlag, 3rd edition, 2008.

[BDHT05] Z.-D. Bai, L. Devroye, H.-K. Hwang, and T.-H. Tsai. Maxima in hypercubes. *Random Struct. Alg.*, 27(3):290–309, 2005.

[BKS01] S. Börzsönyi, D. Kossmann, and K. Stocker. The skyline operator. In *Proc. 17th IEEE Int. Conf. Data Eng.*, pages 421–430, 2001.

[BR10a] I. Bárány and M. Reitzner. On the variance of random polytopes. *Adv. Math.*, 225(4):1986–2001, 2010.

[BR10b] I. Bárány and M. Reitzner. Poisson polytopes. *Annals. Prob.*, 38(4):1507–1531, 2010.

[CF90] B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.

[CHR14] H.-C. Chang, S. Har-Peled, and B. Raichel. From proximity to utility: A Voronoi partition of Pareto optima. *CoRR*, abs/1404.3403, 2014.

[Cla04] K. L. Clarkson. On the expected number of k-sets of coordinate-wise independent points. manuscript, 2004.

[CMS93] K. L. Clarkson, K. Mehlhorn, and R. Seidel. Four results on randomized incremental constructions. *Comput. Geom. Theory Appl.*, 3(4):185–212, 1993.

[CS89] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.

[Fel08] A. Feldman. Welfare economics. In S. Durlauf and L. Blume, editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, 2008.

[GSG07] P. Godfrey, R. Shipley, and J. Gryz. Algorithms and analyses for maximal vector computation. *VLDB J.*, 16(1):5–28, 2007.

[Har11] S. Har-Peled. *Geometric Approximation Algorithms*, Volume 173 of *Mathematical Surveys and Monographs*. Amer. Math. Soc., 2011.

[HR14] S. Har-Peled and B. Raichel. On the expected complexity of randomly weighted Voronoi diagrams. In *Proc. 30th Annu. Sympos. Comput. Geom.* (SoCG), pages 232–241, 2014.

[HTC13] H.-K. Hwang, T.-H. Tsai, and W.-M. Chen. Threshold phenomena in k-dominant skylines of random samples. *SIAM J. Comput.*, 42(2):405–441, 2013.

[KLP75] H. Kung, F. Luccio, and F. Preparata. On finding the maxima of a set of vectors. *J. Assoc. Comput. Mach.*, 22(4):469–476, 1975.

[OSW84] T. Ottmann, E. Soisalon-Soininen, and D. Wood. On the definition and computation of rectlinear convex hulls. *Inf. Sci.*, 33(3):157–171, 1984.

[SA95] M. Sharir and P. K. Agarwal. *Davenport-Schinzel Sequences and Their Geometric Applications*. Cambridge University Press, New York, 1995.

[Sei93] R. Seidel. Backwards analysis of randomized geometric algorithms. In J. Pach, editor, *New Trends in Discrete and Computational Geometry*, volume 10 of *Algorithms and Combinatorics*, pages 37–68. Springer-Verlag, 1993.

[Sha03] M. Sharir. The Clarkson-Shor technique revisited and extended. *Comb., Prob. & Comput.*, 12(2):191–201, 2003.

[SW93] R. Schneider and J. A. Wieacker. Integral geometry. In P. M. Gruber and J. M. Wills, editors, *Handbook of Convex Geometry*, volume B, chapter 5.1, pages 1349–1390. North-Holland, 1993.

[WW93] W. Weil and J. A. Wieacker. Stochastic geometry. In P. M. Gruber and J. M. Wills, editors, *Handbook of Convex Geometry*, volume B, chapter 5.2, pages 1393–1438. North-Holland, 1993.

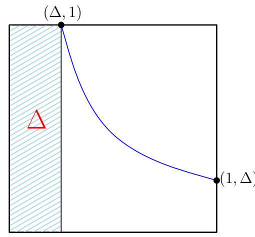


Figure A.1: Graph of the function $xy = c$

A. An Integral Calculation

Lemma A.1. *Let $F_d(\Delta)$ be the total measure of the points $\mathsf{p} = (\mathsf{p}_1, \ldots, \mathsf{p}_d)$ in the hypercube $[0,1]^d$, such that $\mathrm{pv}(\mathsf{p}) = \mathsf{p}_1\mathsf{p}_2 \cdots \mathsf{p}_d \leq \Delta$. That is, $F_d(\Delta)$ is the measure of all points in hypercube with point volume $\leq \Delta$. Then $F_d(\Delta) = \sum_{i=0}^{d-1} \frac{\Delta}{i!} \ln^i \frac{1}{\Delta}$.*

Proof: The claim follows by tedious but relatively standard calculations. As such, the proof is included for the sake of completeness. First consider the simpler $d = 2$ case. Here the points whose point volume equals Δ are defined by the curve $xy = \Delta$. This curve intersects the unit square at the point $(\Delta, 1)$ (see Figure A.1). As $F_d(\Delta)$ is the total volume under this curve in the unit square we have that

$$F_2(\Delta) = \Delta + \int_{x=\Delta}^{1} \frac{\Delta}{x} \mathrm{d}x = \Delta + [\Delta \ln x]_{x=\Delta}^{1} = \Delta + \Delta \ln \frac{1}{\Delta}.$$

Extending the argument one dimension higher we have,

$$\begin{aligned} F_3(\Delta) &= \Delta + \int_{x_3=\Delta}^{1} F_2\left(\frac{\Delta}{x_3}\right) \mathrm{d}x_3 = \Delta + \int_{x_3=\Delta}^{1} \left(\frac{\Delta}{x_3} + \frac{\Delta}{x_3} \ln \frac{x_3}{\Delta}\right) \mathrm{d}x_3 \\ &= \Delta + \Delta \ln \frac{1}{\Delta} + \int_{x_3=\Delta}^{1} \left(\frac{\Delta}{x_3} \ln \frac{x_3}{\Delta}\right) \mathrm{d}x_3 = \Delta + \Delta \ln \frac{1}{\Delta} + \left[\frac{\Delta}{2} \ln^2 \frac{x_3}{\Delta}\right]_{x_3=\Delta}^{1} \\ &= \Delta + \Delta \ln \frac{1}{\Delta} + \frac{\Delta}{2} \ln^2 \frac{1}{\Delta}. \end{aligned}$$

More generally, we have

$$\frac{1}{(d-1)!} \int_{x=\Delta}^{1} \frac{\Delta}{x} \ln^{d-1} \frac{x}{\Delta} \mathrm{d}x = \frac{\Delta}{(d-1)!} \left[\frac{1}{d} \ln^d \frac{x}{\Delta}\right]_{x=\Delta}^{1} = \frac{\Delta}{d!} \ln^d \frac{1}{\Delta}.$$

In particular, assume that

$$F_{d-1}(\Delta) = \sum_{i=0}^{d-2} \frac{1}{i!} \Delta \ln^i \frac{1}{\Delta}.$$

Now, we have that

$$\begin{aligned} F_d(\Delta) &= \Delta + \int_{x_d=\Delta}^{1} F_{d-1}\left(\frac{\Delta}{x_d}\right) \mathrm{d}x_d = \Delta + \int_{x_d=\Delta}^{1} \left(\sum_{i=0}^{d-2} \frac{\Delta}{i! x_d} \ln^i \frac{x_d}{\Delta}\right) \mathrm{d}x_d \\ &= \Delta + \sum_{i=0}^{d-2} \frac{1}{i!} \left(\int_{x_d=\Delta}^{1} \frac{\Delta}{x_d} \ln^i \frac{x_d}{\Delta} \mathrm{d}x_d\right) = \Delta + \sum_{i=1}^{d-1} \frac{\Delta}{i!} \ln^i \frac{1}{\Delta} = \sum_{i=0}^{d-1} \frac{\Delta}{i!} \ln^i \frac{1}{\Delta}. \end{aligned}$$