On the Average Complexity of the k-Level*

Man-Kwun Chiu[1] Stefan Felsner[2] Manfred Scheucher[2]
Patrick Schnider[3] Raphael Steiner[2] Pavel Valtr[4]

[1] Department of Mathematics and Computer Science,
Freie Universität Berlin, Germany,
{chiumk}@zedat.fu-berlin.de

[2] Institut für Mathematik,
Technische Universität Berlin, Germany,
{felsner,scheucher,steiner}@math.tu-berlin.de

[3] Department of Computer Science,
ETH Zürich, Switzerland
{patrick.schnider}@inf.ethz.ch

[4] Department of Applied Mathematics,
Faculty of Mathematics and Physics, Charles University, Czech Republic
{valtr}@kam.mff.cuni.cz

Abstract

Let $\mathcal{L}$ be an arrangement of n lines in the Euclidean plane. The *k-level* of $\mathcal{L}$ consists of all vertices v of the arrangement which have exactly k lines of $\mathcal{L}$ passing below v. The complexity (the maximum size) of the k-level in a line arrangement has been widely studied. In 1998 Dey proved an upper bound of $O(n \cdot (k+1)^{1/3})$. Due to the correspondence between lines in the plane and great-circles on the sphere, the asymptotic bounds carry over to arrangements of great-circles on the sphere, where the k-level denotes the vertices at distance k to a marked cell, the *south pole*.

We prove an upper bound of $O((k+1)^2)$ on the expected complexity of the $(\leq k)$-level in great-circle arrangements if the south pole is chosen uniformly at random among all cells.

We also consider arrangements of great $(d-1)$-spheres on the d-sphere $\mathbb{S}^d$ which are orthogonal to a set of random points on $\mathbb{S}^d$. In this model, we prove that the expected complexity of the k-level is of order $\Theta((k+1)^{d-1})$.

In both scenarios, our bounds are independent of n, showing that the distribution of arrangements under our sampling methods differs significantly from other methods studied in the literature, where the bounds do depend on n.

*M.-K. Chiu was supported by ERC StG 757609. S. Felsner and M. Scheucher were supported by DFG Grant FE 340/12-1. R. M. Scheucher was supported by the internal research funding "Post-Doc-Funding" from Technische Universitt Berlin. Steiner was supported by DFG-GRK 2434. P. Schnider was supported by the SNSF Project 200021E-171681. P. Valtr was supported by the grant no. 18-19158S of the Czech Science Foundation (GAČR) and by the PRIMUS/17/SCI/3 project of Charles University. This work was initiated at a workshop of the collaborative DACH project *Arrangements and Drawings* in Schloss St. Martin, Graz. We thank the organizers for the inspiring atmosphere. We also thank Birgit Vogtenhuber for interesting discussions.

1 Introduction

Let $\mathcal{L}$ be an arrangement of n lines in the Euclidean plane. The *vertices* of $\mathcal{L}$ are the intersection points of lines of $\mathcal{L}$. Throughout this article we consider arrangements to be *simple*, i.e., no three lines intersect in a common vertex. Moreover, we assume that no line is vertical. The *k-level* of $\mathcal{L}$ consists of all vertices v which have exactly k lines of $\mathcal{L}$ below v. The *$(\leq k)$-level* of $\mathcal{L}$ consists of all vertices v which have at most k lines of $\mathcal{L}$ below v. We denote the k-level by $V_k(\mathcal{L})$ and its size by $f_k(\mathcal{L})$. Moreover, by $f_k(n)$ we denote the maximum of $f_k(\mathcal{L})$ over all arrangements $\mathcal{L}$ of n lines, and by $f(n) = f_{\lfloor (n-2)/2 \rfloor}(n)$ the maximum size of the *middle level*.

A *k-set* of a finite point set P in the Euclidean plane is a subset K of k elements of P that can be separated from $P \setminus K$ by a line. Paraboloid duality is a bijection $P \leftrightarrow \mathcal{L}_P$ between point sets and line arrangements (for details on this duality see [O'R94, Chapter 6.5] or [Ede87, Chapter 1.4]). The number of k-sets of P equals $|V_{k-1}(\mathcal{L}_P) \cup V_{n-1-k}(\mathcal{L}_P)|$.

In discrete and computational geometry bounds on the number of k-sets of a planar point set, or equivalently on the size of k-levels of a planar line arrangement have important applications. The complexity of k-levels was first studied by Lovász [Lov71] and Erdős et al. [ELSS73]. They bound the size of the k-level by $O(n \cdot (k+1)^{1/2})$. Dey [Dey98] used the crossing lemma to improve the bound to $O(n \cdot (k+1)^{1/3})$. In particular, the maximum size $f(n)$ of the middle level is $O(n^{4/3})$. Concerning the lower bound on the complexity, Erdős et al. [ELSS73] gave a construction showing that $f(2n) \geq 2f(n)+cn = \Omega(n \log n)$ and conjectured that $f(n) \geq \Omega(n^{1+\varepsilon})$. An alternative $\Omega(n \log n)$-construction was given by Edelsbrunner and Welzl [EW85]. The current best lower bound $f_k(n) \geq n \cdot e^{\Omega(\sqrt{\log k})}$ was obtained by Nivasch [Niv08] improving the constant on a bound of the same asymptotic by Tóth [Tót01]. The complexity of the $(\leq k)$-level in arrangements of lines is better understood. Specifically, Alon and Györi [AG86] prove a tight upper bound of $(k+1)(n-k/2-1)$ for its size. For further information, we recommend the survey by Wagner [Wag07].

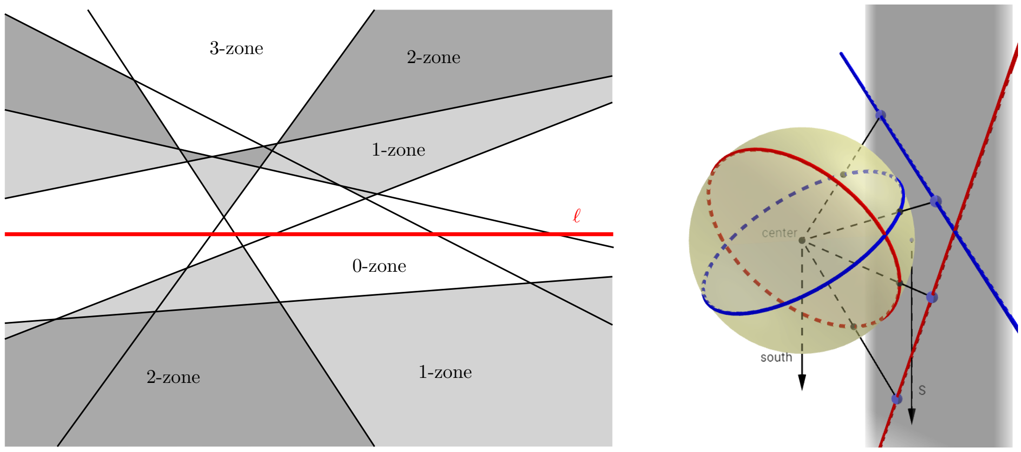
1.1 Generalized Zone Theorem

In order to define "zones", let us introduce the notion of "distances". For x and x' being a vertex, edge, line, or cell of an arrangement $\mathcal{L}$ of lines in $\mathbb{R}^2$ we let their *distance* $\mathrm{dist}_{\mathcal{L}}(x, x')$ be the minimum number of lines of $\mathcal{L}$ intersected by the interior of a curve connecting a point of x with a point of x'. Pause to note that the k-level of $\mathcal{L}$ is precisely the set of vertices which are at distance k to the bottom cell.

The $(\leq j)$-zone $Z_{\leq j}(\ell, \mathcal{L})$ of a line ℓ in an arrangement $\mathcal{L}$ is defined as the set of vertices, edges, and cells from $\mathcal{L}$ which have distance at most j from ℓ. See Figure 1(a) for an illustration.

For arrangements of hyperplanes in $\mathbb{R}^d$ the $(\leq j)$-zone is defined similarly. The classical zone theorem provides bounds for the complexity of the zone ((≤ 0)-zone) of a hyperplane (cf. [ESS91] and [Mat02, Chapter 6.4]). A generalization with bounds for the complexity of the $(\leq j)$-zone appears as an exercise in Matoušek's book [Mat02, Exercise 6.4.2]. In the proof of Theorem 2 we use a variant of the 2-dimensional case (Proposition 1). For the sake of completeness and to provide explicit constants, we include the proof of Proposition 1 in Section 3.

Proposition 1. *Let $\mathcal{L}$ be a simple arrangement of n lines in $\mathbb{R}^2$ and $\ell \in \mathcal{L}$. The $(\leq j)$-zone of ℓ contains at most $2e \cdot (j+1)n$ vertices strictly above ℓ.*



(a) (b)

Figure 1: (a) The higher order zones of a line ℓ. (b) The correspondence between great-circles on the unit sphere and lines in a plane. Using the center of the sphere as the center of projection points on the sphere are projected to the points in the plane.

1.2 Arrangements of Great-Circles

Let Π be a plane in 3-space which does not contain the origin and let $\mathbb{S}^2$ be a sphere in 3-space centered at the origin. The central projection Ψ_Π yields a bijection between arrangements of great circles on $\mathbb{S}^2$ and arrangements of lines in Π. Figure 1(b) gives an illustration.

The correspondence Ψ_Π preserves interesting properties, e.g. simplicity of the arrangements. If $\Psi_\Pi(\mathcal{C}) = \mathcal{L}$ and $\mathcal{L}$ has no parallel lines, then Ψ_Π induces a bijection between pairs of antipodal vertices of $\mathcal{C}$ and vertices of $\mathcal{L}$.

As in the planar case, we define the *distance* between points x, y of $\mathbb{S}^2$ with respect to a great-circle arrangement $\mathcal{C}$ as the minimum number of circles of $\mathcal{C}$ intersected by the interior of a curve connecting x with y. The *k-level* (*$(\leq k)$-level* resp.) of $\mathcal{C}$ is the set of all the vertices of $\mathcal{C}$ at distance k (distance at most k resp.) from the south pole. The $(\leq j)$-zone of a great-circle in $\mathbb{S}^2$ is defined similar to the $(\leq j)$-zone of a line in $\mathbb{R}^2$.

Let Π_1 and Π_2 be two parallel planes in 3-space with the origin between them and let Ψ_1 and Ψ_2 be the respective central projections. For a great-circle arrangement $\mathcal{C}$ we consider $\mathcal{L}_1 = \Psi_1(\mathcal{C})$ and $\mathcal{L}_2 = \Psi_2(\mathcal{C})$. A vertex v from the k-level of $\mathcal{C}$ maps to a vertex of the k-level in one of $\mathcal{L}_1$, $\mathcal{L}_2$ and to a vertex of the $(n-k-2)$-level in the other. Hence, bounds for the maximum size of the k-level of line arrangements carry over to the k-level of great-circle arrangements except for a multiplicative factor of 2.

The $(\leq j)$-zone of a great-circle C in $\mathcal{C}$ projects to a $(\leq j)$-zone of a line in each of $\mathcal{L}_1$ and $\mathcal{L}_2$. Hence, the complexity of a $(\leq j)$-zone in $\mathcal{C}$ is upper bounded by two times the maximum complexity of a $(\leq j)$-zone in a line arrangement. Proposition 1 implies that the $(\leq j)$-zone of a great-circle C in an arrangement of n great-circles contains at most $4e \cdot (j+1)n$ vertices in each of the two open hemispheres bounded by C.

1.3 Higher Dimensions

The problem of determining the complexity of the k-level admits a natural extension to higher dimensions. We consider arrangements in $\mathbb{R}^d$ of hyperplanes to be *simple*, meaning that no $d+1$

hyperplanes intersect in a common point. Moreover, we assume that no hyperplane is parallel to the x_d-axis. The *k-level* of $\mathcal{A}$ consists of all vertices (i.e. intersection points of d hyperplanes) which have exactly k hyperplanes of $\mathcal{A}$ below them (with respect to the d-th coordinate). We denote the k-level by $V_k(\mathcal{A})$ and its size by $f_k(\mathcal{A})$. Moreover, by $f_k^{(d)}(n)$ we denote the maximum of $f_k(\mathcal{A})$ among all arrangements $\mathcal{A}$ of n hyperplanes in $\mathbb{R}^d$.

As in the planar case, there remains a gap between lower and upper bounds;

$$\Omega(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - 1}) \leq f_k^{(d)}(n) \leq O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - c_d}),$$

here $c_d > 0$ is a small positive constant only depending on d. Details and references can be found in Chapter 11 of Matoušek's book [Mat02]. In dimensions 3 and 4 improved bounds have been established. For example, for $d = 3$, it is known that $f_k^{(3)}(n) \leq O(n(k+1)^{3/2})$ (see [SST01]). For the middle level in dimension $d \geq 2$ an improved lower bound $f^{(d)}(n) \geq n^{d-1} \cdot e^{\Omega(\sqrt{\log n})}$ is known (see [Tót01] and [Niv08]).

We call the intersection of $\mathbb{S}^d$ with a central hyperplane in $\mathbb{R}^{d+1}$ a *great-$(d-1)$-sphere* of $\mathbb{S}^d$. Similar to the planar case, arrangements of hyperplanes in $\mathbb{R}^d$ are in correspondence with arrangements of great-$(d-1)$-spheres on the unit sphere $\mathbb{S}^d$ (embedded in $\mathbb{R}^{d+1}$). The terms "distance" and "k-level" generalize in a natural way.

2 Our Results

In the first part of this paper we consider arrangements of great-circles on the sphere and investigate the average complexity of the k-level when the southpole is chosen uniformly at random among the cells. This question was raised by Barba, Pilz, and Schnider while sharing a pizza [BPS19, Question 4.2].

In Section 4 we prove the following bound on the average complexity.

Theorem 2. *Let $\mathcal{C}$ be a simple arrangement of great-circles. The expected size of the $(\leq k)$-level is at most $16e \cdot (k+2)^2$ when the southpole is chosen uniformly at random among the cells of $\mathcal{C}$.*

Note that for $k \geq n/4$ the bound is meaningless, since it exceeds the number of vertices of the arrangement. Our proof works for $k < n/3$ which is needed for Lemma 5. It is remarkable that the bound is independent of the number n of great-circles in the arrangement.

In the second part, we investigate arrangements of randomly chosen great-circles. Here we propose the following model of randomness. On $\mathbb{S}^2$ we have the duality between points and great-circles (each antipodal pair of points defines the normal vector of the plane containing a great-circle). Since we can choose points uniformly at random from $\mathbb{S}^2$, we get random arrangements of great-circles. The duality generalizes to higher dimensions so that we can talk about random arrangements on $\mathbb{S}^d$ for a fixed dimension $d \geq 2$. Using the duality between antipodal pairs of points on $\mathbb{S}^d$ and great-$(d-1)$-spheres, we prove the following bound on the expected size of the k-level in this random model (the proof can be found in Section 5). Again the bound does not depend on the size of the arrangement.

Theorem 3. *Let $d \geq 2$ be fixed. In an arrangement of n great-$(d-1)$-spheres chosen uniformly at random on the unit sphere $\mathbb{S}^d$ (embedded in $\mathbb{R}^{d+1}$), the expected size of the k-level is of order $\Theta((k+1)^{d-1})$ for all $k \leq n/2$.*

3 Proof of Proposition 1

As hinted in Matoušek's book [Mat02, Exercise 6.4.2], we use the method of Clarkson and Shor [CS89] to prove Proposition 1.

Let $\mathcal{L}$ be an arrangement of n lines in $\mathbb{R}^2$ and let $\ell \in \mathcal{L}$ be a fixed line. For any $j = 0, 1, \ldots, n-1$ denote by $V_{\leq j}$ the set of vertices of $\mathcal{L}$ contained in the $(\leq j)$-zone $Z_{\leq j}(\ell, \mathcal{L})$ of ℓ and lying strictly above ℓ. In other words, $v \in V_{\leq j}$ if there is a simple path P_v in the halfplane ℓ^+ from v to ℓ whose interior has at most j intersections with lines from $\mathcal{L}$.

Let R be a random sample of lines from $\mathcal{L}$ where $\ell \in R$ and each line $\ell' \neq \ell$ independently belongs to R with probability $p := \frac{1}{j+1}$. The probability that a vertex $v \in V_{\leq j}$ is present in the induced subarrangement $\mathcal{L}(R)$ and appears at distance 0 from ℓ is at least $(\frac{1}{j+1})^2 \cdot (1 - \frac{1}{j+1})^r$, where $0 \leq r \leq j$ denotes the distance of v from ℓ in $\mathcal{L}$. Figure 2 gives an illustration. Note that

$$\left(1 - \frac{1}{j+1}\right)^r \geq \left(1 - \frac{1}{j+1}\right)^j = \left(\frac{j}{j+1}\right)^j = \left(1 + \frac{1}{j}\right)^{-j} \geq 1/e.$$

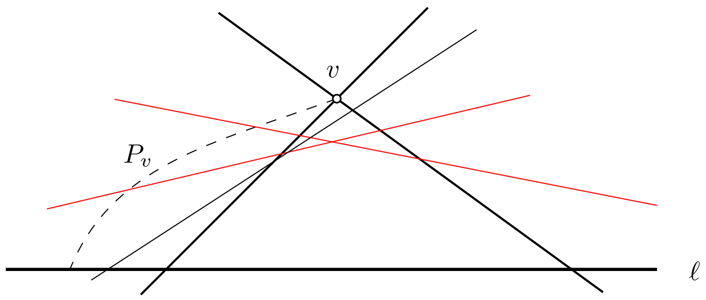


Figure 2: A path P_v witnessing that v belongs to the $(\leq j)$-zone of ℓ for all $j \geq 2$.

Let X be the number of vertices in the 0-zone of ℓ in $\mathcal{L}(R)$ that lie strictly above ℓ. For the expectation of this random variable we have

$$\mathbb{E}(X) \geq \frac{1}{e}\left(\frac{1}{j+1}\right)^2 \cdot |V_{\leq j}|.$$

An inductive argument, as used to show the classical zone theorem (see [GHW13, page 136]), shows there are at most $2n-3$ vertices lying strictly above ℓ in the zone. Hence, we have $X \leq 2 \cdot |R|$ and

$$\mathbb{E}(X) \leq 2 \cdot \mathbb{E}(|R|) = 2np.$$

The above inequalities imply

$$|V_{\leq j}| \leq e \cdot (j+1)^2 \cdot 2 \cdot n \cdot p = 2 \cdot e \cdot (j+1) \cdot n.$$

This concludes the proof of the theorem.

4 Proof of Theorem 2

For the proof of Theorem 2, we fix a great-circle C from $\mathcal{C}$ and denote the two closed hemispheres bounded by C on $\mathbb{S}^2$ as C^+ and C^-. As an intermediate step, we bound the size of the set

$\mathcal{F}_{\le k}(C^+)$ of pairs (F, v), where F is a cell of C^- touching C and v is a vertex of C^+ whose distance to F is at most k. The main ingredient to the proof of the theorem is to show $|\mathcal{F}_{\le k}(C^+)| \le 8e \cdot (k+1)^2 n$. We begin with auxiliary considerations.

Consider a family $\mathcal{I}$ of half-intervals in $\mathbb{R}$, it consists of *left-intervals* of the form $(-\infty, a]$ and *right-intervals* $[b, \infty)$. A subset J of k half-intervals from $\mathcal{I}$ is a *k-clique* if there is a point $p \in \mathbb{R}$ that lies in all the half-intervals of J but not in any half-interval of $\mathcal{I} \setminus J$. Similarly, a *$(\le k)$-clique* is defined as a clique of size at most k.

Lemma 4. *Any family $\mathcal{I}$ of half-intervals in $\mathbb{R}$ contains at most $2k+1$ different $(\le k)$-cliques.*

Proof. For $p \in \mathbb{R}$, let $l(p)$ be the number of left-intervals and $r(p)$ the number of right-intervals containing p. A point p certifies a $(\le k)$-clique if and only if $l(p) + r(p) \le k$. From the monotonicity of the functions l and r it follows that if $(l(p_1), r(p_1)) = (l(p_2), r(p_2))$ for two points p_1 and p_2, then they are contained in the same sub-interval. Thus, they certify the same clique. In other words, when we move from one sub-interval to its right sub-interval, either l is decreased by 1 or r is increased by 1. We proceed to bound the number of sub-intervals corresponding to (l, r)-pairs whose sum is at most k.

Let I_1 be the leftmost sub-interval such that its (l, r)-pair (l_1, r_1) satisfies $l_1 + r_1 \le k$, and let I_2 be the rightmost sub-interval such that its (l, r)-pair (l_2, r_2) satisfies $l_2 + r_2 \le k$. The number of sub-intervals between I_1 and I_2 (including them) is $l_1 - l_2 + r_2 - r_1 + 1$ because of the monotonicity of l- and r-values. This number is at most $2k+1$ because $l_2, r_1 \ge 0$ and $l_1, r_2 \le k$. Now, the definition of I_1 and I_2 implies that the number of $(\le k)$-cliques is most $2k+1$. $\square$

The next lemma is a corresponding result for half-circles on the circle $\mathbb{S}^1$.

Lemma 5. *Any family $\mathcal{H}$ of n half-circles in $\mathbb{S}^1$ with $n > 3k$ contains at most $2k+1$ different $(\le k)$-cliques.*

Proof. For this proof, we embed $\mathbb{S}^1$ as the unit-circle in $\mathbb{R}^2$, which is centered at the origin $\mathbf{o}$. We consider the set X of all points from $\mathbb{S}^1$, which are contained in at most k of the half-circles of $\mathcal{H}$, and distinguish the following two cases.

Case 1: The origin $\mathbf{o}$ is not contained in the convex hull of X. There is a line separating $\mathbf{o}$ from X and rotational symmetry allows us to assume that X is contained in the half-plane $\Pi^+ = \{(x, y) \in \mathbb{R}^2 \colon y > 0\}$. For each half-circle $C \in \mathcal{H}$, the central projection of $C \cap \Pi^+$ to the line $y = 1$ is a half-interval. Since $(\le k)$-cliques of $\mathcal{H}$ and $(\le k)$-cliques of the half-intervals are in bijection we get from Lemma 4 that $\mathcal{H}$ has at most $2k+1$ different $(\le k)$-cliques.

Case 2: The origin $\mathbf{o}$ is contained in the convex hull of X. By Carathéodory's theorem, we can find three points p_1, p_2, p_3 such that $\mathbf{o}$ lies in the convex hull of p_1, p_2, p_3. Since each of the n half-circles from $\mathcal{H}$ contains at least one of these three points, and each of these three points lies on at most k half-circles, we have $n \le 3k$, which contradicts the assumption that $n > 3k$. $\square$

For a fixed vertex $v \in C^+ \setminus C$, let $\mathcal{B}_{C^+}(v)$ be the set of cells F such that $(F, v) \in \mathcal{F}_{\le k}(C^+)$, in particular $\mathrm{dist}(F, v) \le k$.

Claim. $|\mathcal{B}_{C^+}(v)| \le 2k - 1$.

Proof. Consider a great-circle $D \ne C$ from $\mathcal{C}$. For a point $x \in C$, we say that (v, x) is *D-separated* if every path from v to x in C^+ intersects D. The set of all *D-separated* points forms a half-circle H_D on C. Let $\mathcal{H}$ be the set of these half-circles, i.e., $\mathcal{H} = \{H_D : D \in \mathcal{C}, D \ne C\}$. See Figure 3.

We claim that there is a bijection between $\mathcal{B}_{C^+}(v)$ and the $(\le k-1)$-cliques in $\mathcal{H}$. Indeed, if the intersection of the half-circles of a clique K, viewed as a subset of C, is I_K, then I_K is the

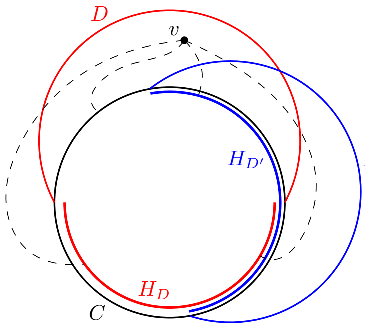


Figure 3: An illustration of the cyclic half-circles $\mathcal{H}$.

interval of C which is reachable from v by crossing the circles corresponding to the half-circles of K. If F is a cell from C^- at distance $i \leq k$ from v, then C and a subset of $i-1$ additional circles have to be crossed to reach v from F, i.e., there is a $(\leq k-1)$-clique in $\mathcal{H}$ whose intersection is $F \cap C$. The number of $(\leq k-1)$-cliques in $\mathcal{H}$ is at most $2k-1$ by Lemma 5. $\square$

Claim. $|\mathcal{F}_{\leq k}(C^+)| \leq 8e \cdot (k+1)^2 n$.

Proof. In the case of $k=0$, vertex v must be one of the $2n-2$ vertices on C and F is one of the two cells of C^- which is adjacent to v. Hence, $|\mathcal{F}_{\leq 0}(C^+)| \leq 4n \leq 8e \cdot (k+1)^2 n$.

Let $k \geq 1$ and note that if $(F, v) \in \mathcal{F}_{\leq k}(C^+)$ then v belongs to the $(\leq k-1)$-zone of C and $F \in \mathcal{B}_{C^+}(v)$. As already noted in Section 1.2, the $(\leq k-1)$-zone of C contains at most $4e \cdot kn$ vertices of $C^+ \setminus C$ and $2n-2$ vertices on C. From the above claim we have $|\mathcal{B}_{C^+}(v)| \leq 2k-1$ for any $v \in C^+ \setminus C$. For the vertices v on C, there are only $2k+2$ cells of C^- touching C with distance at most k to v. Hence we conclude that $|\mathcal{F}_{\leq k}(C^+)| \leq 4e \cdot kn \cdot (2k-1) + (2n-2) \cdot (2k+2) \leq 8e \cdot (k+1)^2 n$. $\square$

Since C was chosen arbitrarily among all great-circles from $\mathcal{C}$ and C^+ was chosen arbitrarily among the two hemispheres of C, the upper bound from the above claim holds for any induced hemisphere of $\mathcal{C}$. For the union $\mathcal{F}_{\leq k}$ of the $\mathcal{F}_{\leq k}(C^+)$ over all the $2n$ choices of the hemisphere C^+, we have

$$|\mathcal{F}_{\leq k}| \leq \sum_{C^+ \text{ hemisphere}} |\mathcal{F}_{\leq k}(C^+)| \leq 16e(k+1)^2 n^2.$$

Proof of Theorem 2. The $(\leq k)$-level with the southpole chosen in cell F consists of the vertices at distance at most k from F. Thus, the expected complexity of the $(\leq k)$-level when choosing F uniformly at random equals $|\mathcal{F}_{\leq k}|$ divided by the number of cells. Since the number of cells in an arrangement of n great-circles is $2\binom{n}{2}+2$ and $|\mathcal{F}_{\leq k}| \leq 16e(k+1)^2 n^2$, we can conclude the statement from

$$\frac{16e \cdot (k+1)^2 \cdot n^2}{2\binom{n}{2}+2} \leq 16e \cdot (k+1)^2 \cdot \frac{n}{n-1} \leq 16e \cdot (k+2)^2 \cdot \underbrace{\frac{k+1}{k+2} \cdot \frac{n}{n-1}}_{\leq 1}. \qquad \square$$

5 Proof of Theorem 3

Let $\mathcal{C}$ be a simple arrangement of n great-$(d-1)$-spheres on the unit sphere $\mathbb{S}^d = \{x \in \mathbb{R}^{d+1} : \|x\| = 1\}$ with center $\mathbf{o} = (0, \ldots, 0)$ in $\mathbb{R}^{d+1}$. For a vertex v of the arrangement, let $\phi_{\mathcal{C}}(v)$

denote the number of great-$(d-1)$-spheres of $\mathcal{C}$ that are crossed by the geodesic arc from v to the south-pole $\mathbf{s} = (0, \ldots, 0, -1)$ of the sphere. The set of vertices v of $\mathcal{C}$ with $\phi_{\mathcal{C}}(v) = k$ is denoted $V_k(\mathcal{C})$.

When $\mathcal{C}$ is projected to a d-dimensional plane H with the origin $\mathbf{o}$ as center of projection, we obtain an arrangement $\mathcal{A}$ of hyperplanes in $\mathbb{R}^d$. Moreover, if the south pole $\mathbf{s}$ is projected to a point "at infinity" of H, say to $(0, \ldots, 0, -\infty)$, then, for every point p in $\mathbb{S}^d$, the circle in $\mathbb{S}^d$ containing the geodesic arc from p to $\mathbf{s}$ is projected to the "vertical" line through p, i.e., the line $p + (0, \ldots, 0, \lambda)$. The geodesic is projected to one of the two rays starting from p on this line. In particular, all vertices v of $\mathcal{C}$ with $\phi_{\mathcal{C}}(v) = k$ are projected to vertices of $\mathcal{A}$ either at level k or $n - k - d$.

Let $\mathcal{C}$ be an arrangement of randomly chosen great-$(d-1)$-spheres and let $\mathcal{B}$ be a subset of size d in $\mathcal{C}$. Note that with probability 1, the random great-sphere-arrangement is simple, i.e., no great-sphere contains the south-pole and no more than d great-spheres intersect in a common point. Choose p' as one of the two intersection points of the great-$(d-1)$-spheres in $\mathcal{B}$. Now consider the arrangement $\mathcal{C}' = \mathcal{C} - \mathcal{B}$ and note that $(\mathcal{C}', p')$ can be viewed as a random arrangement of great-$(d-1)$-spheres together with a random point on $\mathbb{S}^d$. Hence, to estimate the expected size of $V_k(\mathcal{C})$, we can estimate the probability that $\phi_{\mathcal{C}'}(p') = k$. This is the purpose of the following lemma.

Lemma 6. *Let $\mathcal{C}$ be an arrangement of n great-$(d-1)$-spheres chosen uniformly at random on the unit sphere $\mathbb{S}^d$ (embedded in $\mathbb{R}^{d+1}$ and centered at the origin). Let p be an additional point chosen uniformly at random from $\mathbb{S}^d$, and let A be the geodesic arc from p to the south pole on $\mathbb{S}^d$. For all $k \leq n/2$, the probability q_k that exactly k great-$(d-1)$-spheres from $\mathcal{C}$ intersect A is in $\Theta((k+1)^{d-1}/n^d)$. More precisely, it satisfies*

$$\frac{2^{d-1}\rho\pi(k+1)^{\overline{d-1}}(n-k+1)^{\overline{d-1}}}{(n+1)^{\overline{2d-1}}} \leq q_k \leq \min\left\{\frac{\rho\pi}{n+1}, \frac{\rho\pi^d(k+1)^{\overline{d-1}}}{(n+1)^{\overline{d}}}\right\},$$

where $a^{\overline{b}} = a(a+1)\cdots(a+b-1)$ denotes the rising factorial and $\rho = \rho_d = \frac{\mathrm{area}_{d-1}(\mathbb{S}^{d-1})}{\mathrm{area}_d(\mathbb{S}^d)} = \frac{\Gamma(\frac{d+1}{2})}{\pi^{1/2}\Gamma(\frac{d}{2})}$ only depends on the dimension d.

Proof. Denote by ϕ the length of the geodesic arc A on $\mathbb{S}^d$ from p to $\mathbf{s}$, i.e., ϕ is the angle between the two rays emanating from $\mathbf{o}$ towards $\mathbf{s}$ and p. Note that – independent from the dimension d – the three points $\mathbf{o}$, $\mathbf{s}$, and p lie in a 2-dimensional plane which also contains the geodesic arc A.

Point p lies on a $(d-1)$-sphere C of radius $\sin(\phi)$ in the d-dimensional hyperplane defined by the equation $x_d = -\cos(\phi)$. Figure 4 gives an illustration for the case $d = 2$, where C is a circle.

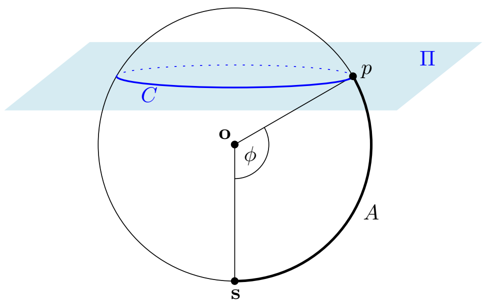


Figure 4: Illustrating the definitions of A, C, and Π depending on p.

The probability that a random great-$(d-1)$-sphere D intersects the arc A defined by the random point p is ϕ/π, since D will intersect the great circle containing A in a random pair of antipodal points. Thus, the probability that A is intersected by exactly k great-$(d-1)$-spheres from the random arrangement $\mathcal{C}$ is

$$q_k = \int_{\phi=0}^{\pi} \underbrace{\frac{\text{area}_{d-1}(\mathbb{S}^{d-1}) \sin^{d-1}(\phi)}{\text{area}_d(\mathbb{S}^d)}}_{\text{density at angle } \phi} \cdot \underbrace{\binom{n}{k} (\phi/\pi)^k (1-\phi/\pi)^{n-k}}_{\text{chosen great-}(d-1)\text{-spheres intersect } A} d\phi.$$

This can be rewritten as

$$q_k = \rho \cdot \binom{n}{k} \cdot \int_{\phi=0}^{\pi} \sin^{d-1}(\phi) \cdot (\phi/\pi)^k (1-\phi/\pi)^{n-k} d\phi,$$

where $\rho = \rho(d) = \frac{\text{area}_{d-1}(\mathbb{S}^{d-1})}{\text{area}_d(\mathbb{S}^d)} = \frac{\Gamma(\frac{d+1}{2})}{\pi^{1/2}\Gamma(\frac{d}{2})}$ is a constant only depending on d. The latter equation follows from $\text{area}_d(\mathbb{S}^d) = 2\pi^{\frac{d+1}{2}}/\Gamma(\frac{d+1}{2})$, where $\Gamma(x)$ is the Euler gamma function (see e.g. [Wikb]).

In the following we give upper and lower bounds for q_k. The Euler beta function B turns out to be the tool to evaluate the integrals:

$$B(a+1, b+1) = \int_{t=0}^{1} t^a (1-t)^b dt = \frac{a! \cdot b!}{(a+b+1)!}.$$

For this identity and more information see for example [Wika].

To show the first upper bound on q_k, we bound the integral above as follows: Since $\sin(\phi) \leq 1$ holds for every $\phi \in [0, \pi]$, we have

$$\begin{aligned} q_k &\leq \rho \binom{n}{k} \int_{\phi=0}^{\pi} (\phi/\pi)^k (1-\phi/\pi)^{n-k} d\phi = \rho\pi \binom{n}{k} \int_{t=0}^{1} t^k (1-t)^{n-k} dt \\ &= \rho\pi \binom{n}{k} B(k+1, n-k+1) = \rho\pi \cdot \frac{n!}{k!(n-k)!} \cdot \frac{k!(n-k)!}{(n+1)!} = \rho\pi \cdot \frac{1}{n+1}. \end{aligned}$$

Towards the second upper bound on q_k, we use the fact that $\sin(\phi) \leq \phi$ holds for every $\phi \in [0, \pi]$:

$$\begin{aligned} q_k &\leq \rho\pi^{d-1} \binom{n}{k} \int_{\phi=0}^{\pi} (\phi/\pi)^{k+d-1} (1-\phi/\pi)^{n-k} d\phi \\ &= \rho\pi^d \binom{n}{k} \int_{t=0}^{1} t^{k+d-1} (1-t)^{n-k} dt \\ &= \rho\pi^d \cdot \frac{n!}{k!(n-k)!} \cdot \frac{(k+d-1)!(n-k)!}{(n+d)!} = \rho\pi^d \cdot \frac{(k+1)^{\overline{d-1}}}{(n+1)^{\overline{d}}}. \end{aligned}$$

To show the lower bound on q_k, we split the integral in two parts: Since $\sin(\phi) \geq 2 \cdot \frac{\phi}{\pi}$ holds for every $\phi \in [0, \pi/2]$ and $\sin(\phi) \geq 2 \cdot (1 - \frac{\phi}{\pi})$ holds for every $\phi \in [\pi/2, \pi]$, we have

$$\begin{aligned}
q_k &\geq 2^{d-1}\rho\binom{n}{k}\left[\int_{\phi=0}^{\pi/2}(\phi/\pi)^{k+d-1}(1-\phi/\pi)^{n-k}d\phi + \int_{\phi=\pi/2}^{\pi}(\phi/\pi)^{k}(1-\phi/\pi)^{n-k+d-1}d\phi\right] \\
&\geq 2^{d-1}\rho\binom{n}{k}\int_{\phi=0}^{\pi}(\phi/\pi)^{k+d-1}(1-\phi/\pi)^{n-k+d-1}d\phi \\
&= 2^{d-1}\rho\pi\binom{n}{k}\int_{t=0}^{1}t^{k+d-1}(1-t)^{n-k+d-1}dt \\
&= 2^{d-1}\rho\pi \cdot \frac{n!}{k!(n-k)!} \cdot \frac{(k+d-1)!(n-k+d-1)!}{(n+2d-1)!} \\
&= \frac{2^{d-1}\rho\pi(k+1)^{\overline{d-1}}(n-k+1)^{\overline{d-1}}}{(n+1)^{\overline{2d-1}}}.
\end{aligned}$$

This completes the proof of Lemma 6. □

Proof of Theorem 3. Consider an arrangement $\mathcal{C}$ of $n+d$ great-$(d-1)$-spheres $C_1, \ldots, C_{n+d}$ chosen uniformly and independently at random from $\mathbb{S}^d$. Let p be a vertex of $\mathcal{C}$ chosen uniformly at random from the intersection points of $\mathcal{C}$ (i.e., one of the two points of intersection of d great-$(d-1)$-spheres $C_{i_1}, \ldots, C_{i_d}$ chosen u.a.r. from $\mathcal{C}$). Note that p is a u.a.r. chosen point from $\mathbb{S}^d$.

We now apply Lemma 6 with p and $\mathcal{C}_p := \mathcal{C} - \{C_{i_1}, \ldots, C_{i_d}\}$. Point p is separated from $\mathbf{s}$ by k great-$(d-1)$-spheres from $\mathcal{C}_p$ with probability $q_k = \Theta(k^{d-1}/n^d)$. Since p is chosen uniformly at random among the $2\binom{n+d}{d}$ vertices of $\mathcal{C}$, we obtain the desired bound of $\Theta(k^{d-1})$ for the number of vertices at distance k from $\mathbf{s}$. □

6 Discussion

Theorem 2 is about arrangements of great-circles. All the elements of the proof, however, carry over to great-pseudocircles whence the result could also be stated for arrangements of great-pseudocircles. Projective arrangements of lines are obtained by antipodal identification from arrangements of great-circles. Hence, if you pick a cell u.a.r. in a projective arrangement of lines (pseudo-lines) the expected number of vertices at distance at most k from the cell is as in Theorem 2. If the projection Ψ_Π is used to project an arrangements $\mathcal{C}$ of great-pseudocircles to an Euclidean arrangement $\mathcal{L}$ on Π such that the south-poles coincide, then the k-level of $\mathcal{C}$ corresponds to the union of the k- and the $(n-k-2)$-level of $\mathcal{L}$.

With respect to lower bounds we would like to know the answer to:

Question 1. *Is there a family of arrangements where the expected size of the middle level is superlinear when the southpole is chosen uniformly at random?*

Recursive constructions from [EW85] and [ELSS73] show that the size of the $(n/2-s)$-level can be in $\Omega(n \log n)$ for any fixed s. Nevertheless computer experiments suggest that if we choose a random southpole for these examples the expected size of the middle level drops to be linear.

Theorem 3 deals with the average size of the k-level in arrangements of randomly chosen great-circles. In our model, great-circles are chosen independently and uniformly at random from the sphere. Since point sets, line arrangements, and great-circle arrangements are in strong correspondence, the bound from Theorem 3 also applies to k-sets in point sets and k-levels of line arrangements from a specific random distribution.

In the context of Erdős–Szekeres-type problems, several articles made use of point sets which are sampled uniformly at random from a convex shape K [BF87, Val95, BGAS13, BSV20]. The average size of the convex hull (0-level) is well-studied for such sets of points. If K is a disk, the convex hull has expected size $O(n^{1/3})$, and if K is a convex polygon with m sides, the expected size is $O(m \log n)$ [HP11, PS85, Ray70, RS63]. Bárány and Steiger have studied the expected number of all k-sets for point sets that are sampled uniformly at random from a convex shape and other random point sets, such as a spherically symmetric distribution in $\mathbb{R}^d$ [BS94]. However, all their resulting bounds depend on n. In particular, the expected size of the convex hull is not constant, which is a substantial contrast to our setting. In fact, our setting appears to be closer to the setting of random order types, for which the expected size of the convex hull was recently shown to be $4 + o(1)$ [GW20]. Hence it would be very interesting to obtain bounds on the average number of k-sets also in this setting. Last but not least, Edelman [Ede92] showed that the expected number of k-sets of an allowable sequence is of order $\Theta(\sqrt{kn})$.

References

[AG86] N. Alon and E. Győry. The number of small semispaces of a finite set of points in the plane. *Journal of Combinatorial Theory, Series A*, 41:154–157, 1986.

[BF87] I. Bárány and Z. Füredi. Empty simplices in Euclidean space. *Canadian Mathematical Bulletin*, 30(4):436–445, 1987.

[BGAS13] J. Balogh, H. Gonzlez-Aguilar, and G. Salazar. Large convex holes in random point sets. *Computational Geometry*, 46(6):725–733, 2013.

[BPS19] L. Barba, A. Pilz, and P. Schnider. Sharing a pizza: bisecting masses with two cuts, 2019. arXiv:1904.02502.

[BS94] I. Bárány and W. Steiger. On the expected number of k-sets. *Discrete & Computational Geometry*, 11(3):243–263, 1994.

[BSV20] M. Balko, M. Scheucher, and P. Valtr. Holes and islands in random point sets, 2020. To appear at the 36th International Symposium on Computational Geometry (SoCG 2020).

[CS89] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete & Computational Geometry*, 4(5):387–421, 1989.

[Dey98] T. K. Dey. Improved bounds for planar k-sets and related problems. *Discrete & Computational Geometry*, 19(3):373–382, 1998.

[Ede87] H. Edelsbrunner. *Algorithms in Combinatorial Geometry.* Springer, 1987.

[Ede92] P. H. Edelman. On the average number of k-sets. *Discrete & Computational Geometry*, 8:209–213, 1992.

[ELSS73] P. Erdős, L. Lovász, G. J. Simmons, and E. G. Straus. Chapter 13 - dissection graphs of planar point sets. In *A Survey of Combinatorial Theory*, pages 139–149. North-Holland, 1973.

[ESS91] H. Edelsbrunner, R. Seidel, and M. Sharir. On the Zone Theorem for Hyperplane Arrangements. In *New Results and New Trends in Computer Science*, pages 108–123. Springer, 1991.

[EW85] H. Edelsbrunner and E. Welzl. On the number of line separations of a finite set in the plane. *Journal of Combinatorial Theory, Series A*, 38:15–29, 1985.

[GHW13] B. Gärtner, M. Hoffmann, and E. Welzl. Computational geometry – lecture notes, 2013. `https://www.ti.inf.ethz.ch/ew/Lehre/CG13/lecture/cg-2013.pdf`.

[GW20] X. Goaoc and E. Welzl. Convex hulls of random order types. In *36th International Symposium on Computational Geometry (SoCG 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.

[HP11] S. Har-Peled. On the expected complexity of random convex hulls, 2011. arXiv:1111.5340.

[Lov71] L. Lovász. On the number of halving lines. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae Sectio Mathematica*, 14:107–108, 1971.

[Mat02] J. Matoušek. *Lectures on Discrete Geometry*. Springer, 2002.

[Niv08] G. Nivasch. An improved, simple construction of many halving edges. *Contemporary Mathematics*, 453:299–305, 2008.

[O'R94] J. O'Rourke. *Computational Geometry in C*. Cambridge University Press, 1994.

[PS85] F. P. Preparata and M. I. Shamos. *Computational Geometry: An Introduction*. Springer, 1985.

[Ray70] H. Raynaud. Sur l'enveloppe convexe des nuages de points aleatoires dans $\mathbb{R}^n$. *Journal of Applied Probability*, 7(1):35–48, 1970.

[RS63] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 2(1):75–84, 1963.

[SST01] M. Sharir, S. Smorodinsky, and G. Tardos. An improved bound for k-sets in three dimensions. *Discrete & Computational Geometry*, 26(2):195–204, Jan 2001.

[Tót01] G. Tóth. Point Sets with Many k-Sets. *Discrete & Computational Geometry*, 26(2):187–194, 2001.

[Val95] P. Valtr. On the minimum number of empty polygons in planar point sets. *Studia Scientiarum Mathematicarum Hungarica*, pages 155–163, 1995.

[Wag07] U. Wagner. k-sets and k-facets. In *Contemporary Mathematics*. Citeseer, 2007.

[Wika] Wikipedia. Beta function. `http://en.wikipedia.org/wiki/Beta_function`.

[Wikb] Wikipedia. N-sphere. `http://en.wikipedia.org/wiki/N-sphere`.