The Number of Combinatorially Different Convex Hulls of Points in Lines*

Heuna Kim† Wolfgang Mulzer† Eunjin Oh‡

Abstract

Given a sequence $\mathcal{R}$ of planar lines in general position, we can obtain a point set by picking exactly one point from each line in $\mathcal{R}$. We provide exponential upper and lower bounds on the number of combinatorially different convex hulls for point sets that are generated in this manner.

1 Introduction

Recently, geometric problems arising from data imprecision have been studied intensively (e.g., [5]). One popular model is as follows: given a sequence $\mathcal{R} = \langle R_1, \dots, R_n \rangle$ of planar regions, we are promised that our eventual input P contains exactly one point from each R_i. The question is whether the knowledge of $\mathcal{R}$ helps to speed up computations on P. There are many algorithmic results for this setting [1, 2, 4, 6]. Here, we are interested in the *combinatorial* question arising from this model: how much does the knowledge of $\mathcal{R}$ restrict the combinatorial structure of P? We consider the case that (1) $\mathcal{R}$ consists of planar lines, and (2) we measure "variance" by the number of combinatorially different convex hulls. It is known that this setting can be exploited algorithmically [3], and we provide exponential upper and lower bounds on the number of possible convex hulls.

Problem Statement. Let $P = \langle p_1, \dots, p_n \rangle$ be a sequence of n points in the plane, and let $\text{conv}(P)$ be the convex hull of P. Let $\langle p_{i_1}, p_{i_2}, \dots, p_{i_k} \rangle$ be the vertices of $\text{conv}(P)$ in clockwise order, such that $i_1 = \min\{i_1, \dots, i_k\}$. We define the *hull signature* of P as $\sigma(P) = \langle i_1, \dots, i_k \rangle$.

Consider a sequence $\mathcal{R} = \langle \ell_1, \dots, \ell_n \rangle$ of n planar lines in *general position*, i.e., every two lines in $\mathcal{R}$ intersect in exactly one point and no three lines in $\mathcal{R}$ have a common intersection. A sequence $P = \langle p_1, \dots, p_n \rangle$ of points in the plane is *restricted to* $\mathcal{R}$ if $p_i \in \ell_i$, for $i = 1, \dots, n$. Given $\mathcal{R}$, we would like to study the set $C(\mathcal{R}) = \{\sigma(P) \mid P \text{ is restricted to } \mathcal{R}\}$ of all hull signatures that can be generated by point sequences restricted to $\mathcal{R}$. We also define $c(\mathcal{R}) = |C(\mathcal{R})|$ and $c(n) = \max_{\mathcal{R}} c(\mathcal{R})$, where $\mathcal{R}$ ranges over all sequences of n planar lines in general position. We provide upper and lower bounds on $c(n)$.

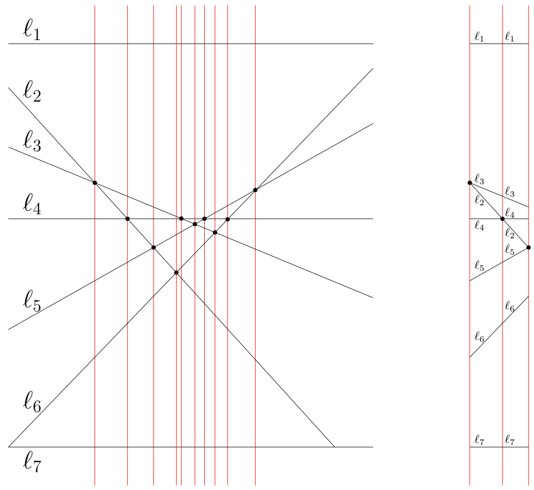

Figure 1: (left) A lower bound example with seven lines. The red lines illustrate the subdivision of the plane into $O(n^2)$ slabs. (right) Between the second and the third slab, ℓ_2 and ℓ_4 change position.

2 Lower Bound

Theorem 1 *For any $n \geq 3$, there exists a sequence $\mathcal{R}$ of n lines in the plane with $c(\mathcal{R}) = \Omega(n^2 3^n)$.*

Proof. Let $\langle \ell_2, \ell_3, \dots, \ell_{n-1} \rangle$ be a sequence of $n-2$ planar lines in general position, and let $Q = \{\ell_i \cap \ell_j \mid 2 \leq i < j \leq n-1\}$ be their intersection points. Set $m = |Q|$. By general position, we have $m = \binom{n-2}{2}$. We endow the plane with a Cartesian coordinate system such that the points in Q have pairwise distinct x-coordinates. We pick two additional lines ℓ_1 and ℓ_n so that all points in Q lie below ℓ_1 and above ℓ_n along the y-axis, and we set $\mathcal{R} = \langle \ell_1, \ell_2 \dots, \ell_{n-1}, \ell_n \rangle$; see Fig. 1. We will show that $c(\mathcal{R}) = \Omega(n^2 3^n)$.

For this, we subdivide the plane into *vertical slabs* as follows: let $a_1, \dots, a_m$ be the sorted sequence of x-coordinates of the points in Q. For $j = 1, \dots, m-1$, we define the *jth vertical slab* $V_j = \{(x, y) \in \mathbb{R}^2 \mid a_j < x < a_{j+1}\}$. By general position, every vertical slab intersects all lines in $\mathcal{R}$, and by definition, no slab contains an intersection point from Q. Furthermore,

*This work was partially supported by research grant AL 253/8-1 from Deutsche Forschungsgemeinschaft (German Science Association).

†Institut für Informatik, Freie Universität Berlin, Germany. heunak@mi.fu-berlin.de,mulzer@inf.fu-berlin.de

‡Dept. Computer Science and Engineering, Pohang University of Science and Technology jin9082@postech.ac.kr

in every vertical slab the line ℓ_1 lies above all other lines and the line ℓ_n lies below all other lines.

Now, we consider for each slab V_j the point sequences that are restricted to the line segments in V_j. More precisely, define $\mathcal{R}_j = \langle \ell_{1j}, \dots, \ell_{nj} \rangle$ as the sequence of line segments $\ell_{ij} = \ell_i \cap V_j$, for $i = 1, \dots n$. We have $C(\mathcal{R}_k) \subseteq C(\mathcal{R})$. Next, we will provide a lower bound on how many distinct hull signatures $C(\mathcal{R}_j)$ contributes to $C(\mathcal{R})$.

We begin with $C(\mathcal{R}_1)$. Fix two arbitrary points $p_1 \in \ell_{11}$ and $p_n \in \ell_{n1}$. The line segment p_1p_n lies inside V_1 and intersects all line segments ℓ_{i1}, for $i = 2, \dots n-1$. By construction, p_1 and p_n always appear on conv(P), no matter how we pick $p_2, \dots, p_{n-1}$ restricted to $\ell_{21}, \dots, \ell_{(n-1)1}$. For every other point p_i, we have three choices: we can pick $p_i \in \ell_{i1}$ such that (i) the index i does not occur in the hull signature $\sigma(P)$ (i.e., p_i is not on conv(P)); (ii) the index i occurs in $\sigma(P)$ before n; or (iii) the index i occurs in $\sigma(P)$ after n. The choice is not completely free: to avoid degeneracies, conv(P) must have at least three vertices, so at least one other point needs to appear in $\sigma(P)$. Thus, the number of distinct hull signatures for point sequences restricted to $\mathcal{R}_1$ is at least $3^{n-2} - 1$.

Now fix $j \in \{2, \dots, m-1\}$ and consider $C(\mathcal{R}_j)$. We give a lower bound on $\left| C(\mathcal{R}_j) \setminus \bigcup_{k=1}^{j-1} C(\mathcal{R}_k) \right|$, the number of hull signatures generated by point sequences restricted to $\mathcal{R}_j$ that are not generated by a point sequence restricted to a previous slab. By construction, the left boundary of $\overline{V_j}$ contains exactly one point $q \in Q$. Let q be the intersection of two lines $\ell_a, \ell_b \in \mathcal{R}$, such that ℓ_a is above ℓ_b to the left of q and ℓ_a is below ℓ_b to the right of q. We pick two arbitrary points $p_1 \in \ell_{1j}$ and $p_n \in \ell_{nj}$. As before, p_1 and p_n always appear on conv(P), the line segment p_1p_n is contained in V_j, and all other line segments ℓ_{ij} intersect p_1p_n, for $i = 2, \dots, n-1$. Next, we take points $p_a \in \ell_{aj}$ and $p_b \in \ell_{bj}$ to the left of p_1p_n. As long as p_a and p_b lie on conv(P), the signature $\sigma(P)$ cannot appear in $\bigcup_{k=1}^{j-1} C(\mathcal{R}_k)$, since in all previous slabs we have that if a and b both occur after n in $\sigma(P)$, then b must come before a. In $C(\mathcal{R}_k)$, however, a comes before b in this case, since ℓ_a and ℓ_b have switched. For all other points p_i, there are again three choices: the index i may appear before or after n in $\sigma(P)$, or it may be absent from $\sigma(P)$. Thus, the slab V_j contributes at least 3^{n-4} new signatures to $\bigcup_{k=1}^{j} C(\mathcal{R}_k)$. There are $\Omega(n^2)$ slabs, so $c(\mathcal{R}) = \Omega(n^2 3^n)$. $\square$

3 Upper Bound

Let $\mathcal{R} = \langle \ell_1, \dots, \ell_n \rangle$ be a sequence of n planar lines in general position. We denote by $A(\mathcal{R})$ the *arrangement* of $\mathcal{R}$, i.e., the subdivision of the plane into *cells*, *edges* and *vertices* induced by the lines in $\mathcal{R}$: a *cell* is a maximal connected component of $\mathbb{R}^2 \setminus \bigcup_{\ell \in \mathcal{R}} \ell$; an *edge* is a maximal component of a line in $\mathcal{R}$ that does not belong to any other line in $\mathcal{R}$; and a *vertex* is the intersection of two lines in $\mathcal{R}$.

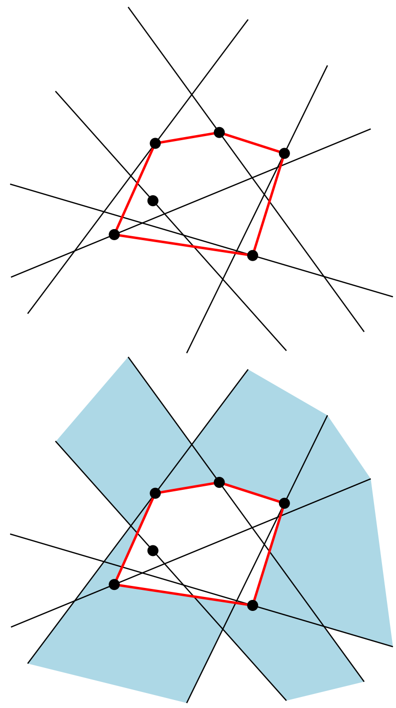

Figure 2: (top) A set $\mathcal{R}$ of lines, a point set P restricted to $\mathcal{R}$, and conv(P). (bottom) The outer zone of conv(P). There are 11 outer regions, and the complexity of the outer zone is 39.

Let $P = \langle p_1, \dots, p_n \rangle$ be a sequence of n points restricted to $\mathcal{R}$. We may assume that no point in P coincides with a vertex of $A(\mathcal{R})$ (otherwise we can perturb P slightly). To distinguish them from edges of $A(\mathcal{R})$, we call the edges of conv(P) *arcs*. Let $\langle e_1, \dots, e_k \rangle$ be the arcs of conv(P), in clockwise order, where e_1 comes after the leftmost point of conv(P) in clockwise order. For an arc e_i and a cell F of $A(\mathcal{R})$, we say that e_i *properly crosses* F if the relative interior of e_i has nonempty intersection with F and that e_i *touches* F if $e_i \cap \overline{F} = v$, where $v = \overline{e_i} \cap \overline{e_{i+1}}$. The arc e_i *crosses* F if e_i properly crosses F or touches F. Let $\mathcal{F} = \langle F_1, \dots, F_a \rangle$ be the sequence of cells in $A(\mathcal{R})$ that are crossed by $\langle e_1, \dots, e_k \rangle$, in clockwise order: first the cells crossed by e_1, then the cells crossed by e_2, etc. The same cell may occur several times in $\mathcal{F}$, but each occurrence is due to one *crossing edge* e_i.

Since the vertices of conv(P) lie on the edges of $A(\mathcal{R})$, each cell $F_j \in \mathcal{F}$ is divided into at most two parts by the corresponding crossing edge e_i. We define the *outer region* Z_j as the component of $F_j \setminus e_i$ whose interior does not intersect conv(P). If F_j is touched by e_i, we call Z_j a *touched outer region* and F_j a *touched cell*. In this case, we have $Z_j = F_j$. The sequence $\mathcal{Z} = \langle Z_1, \dots, Z_a \rangle$ is called the *outer zone* of conv(P) in $A(\mathcal{R})$. Each outer region $Z \in \mathcal{Z}$ is a (possibly unbounded) convex polygon. For a non-touched outer region Z, exactly one edge of Z belongs to the boundary of conv(P). We call it the *supporting arc* of Z. All other edges of Z are (possibly unbounded) parts of edges of $A(\mathcal{R})$. For a touched outer region Z, no edge is part of the boundary of *conv*(P). Instead, exactly one edge e of Z is intersected by conv(P). We split e into two subedges, each being a maximal connected component of $e \setminus \text{conv}(P)$. The edges of Z consist of the two subedges of e and the other edges of $A(\mathcal{R})$ incident to Z. We call $e \cap \text{conv}(P)$ the *supporting vertex* of Z. A *vertex* of a (touched or non-touched) outer region Z is a vertex of $A(\mathcal{R})$ incident to Z. In particular, the supporting vertex or the endpoints of the supporting arc are not vertices of Z.

The *complexity* of an outer region Z is defined as the number of edges of Z_j other than the supporting arc. The *complexity of the outer zone* $\mathcal{Z}$ is the sum of the complexities of the outer regions, see Fig. 2.

Lemma 2 *The outer zone complexity is at most* $8n$*.*

Proof. Let ℓ_i be a line in $\mathcal{R}$. By construction, conv(P) intersects ℓ_i, so $\ell_i \setminus \text{conv}(P)$ consists of exactly two unbounded connected components, the two *rays corresponding to* ℓ_i. We orient these rays towards infinity, and we denote by $R = \langle r_1, \ldots, r_{2n} \rangle$ the sequence of all rays that correspond to lines in $\mathcal{R}$.

Consider an outer region Z. Every edge e of Z other than the supporting arc is part of a single ray from R, and we orient e in the same direction as the underlying ray. Now every vertex of Z has either (i) two outgoing incident edges (*out-out* vertex); (ii) two incoming incident edges (*in-in* vertex); or (iii) one incoming and one outgoing incident edge (*in-out* vertex).

Lemma 3 *Let Z be an outer region. Then Z contains no out-out vertex and at most one in-in vertex. If Z has an in-in vertex, then Z is bounded.*

Proof. Suppose that Z contains an out-out vertex v, and let r_1 and r_2 be the two rays with $v = r_1 \cap r_2$. Since Z lies in a cell of $A(\mathcal{R})$, Z is completely contained in the wedge W bounded by the subrays of r_1 and r_2 beginning in v. However, conv(P) is incident to the start vertices of r_1 and r_2, so W does not contain conv(P). It follows that Z cannot be part of the outer zone, a contradiction; see Fig. 3(top).

Next, suppose that Z contains two in-in vertices v_1 and v_2. Suppose further that v_1 comes before v_2 in clockwise order along Z after the supporting arc. All edges between v_1 and v_2 (in clockwise order) are oriented, and both v_1 and v_2 are in-in, so there is at least one out-out vertex between v_1 and v_2; see Fig. 3 (middle). We have just seen that this is impossible.

Finally, suppose that Z contains an in-in vertex v, and let r_1 and r_2 be the two rays with $v = r_1 \cap r_2$.

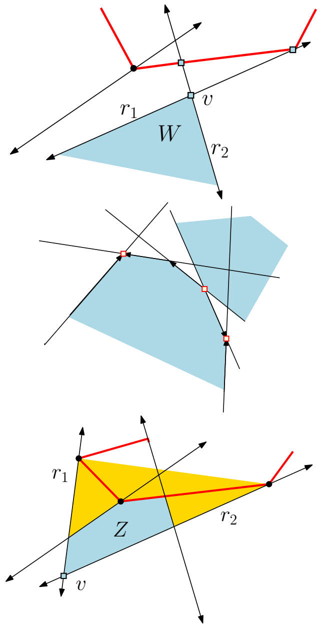

Figure 3: (top) A wedge W bounded by an out-out-vertex v does not contain conv(P). (middle) There cannot be two in-in vertices. (bottom) The outer region Z that contains an in-in vertex v is bounded.

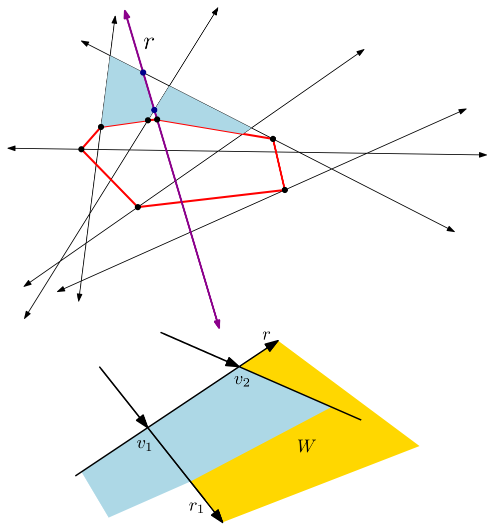

Figure 4: (top) The ray r is charged twice. (bottom) There cannot be two in-out vertices whose outgoing edges lie on the same side of a ray r.

Then Z is completely contained in the region that is bounded by the subrays of r_1 and r_2 from their respective starting points to v, and the boundary of conv(P) between those starting points. It follows that Z is bounded; see Fig. 3(bottom). □

Lemma 4 *There are at most* $4n$ *in-out vertices.*

Proof. Let v be an in-out vertex and let r be the ray that supports the incoming edge of v. We charge v

to r, and we claim that every ray is charged at most twice in this manner; see Fig. 4(top). Indeed, suppose there is a ray r that supports incoming edges for three in-out vertices. Then r contains two in-out vertices v_1 and v_2 such that the outgoing edges for v_1 and v_2 lie on the same side of r. Suppose that v_1 comes before v_2 along r, and let r_1 be the ray supporting the outgoing edge from v_1. Since the rays are directed towards infinity, r and r_1 bound an infinite wedge W with apex v_1. Furthermore, the interior of W is disjoint from conv(P) and thus does not contain any outer region. This contradicts the assumption that the outgoing edge from v_2 extends into W; see Fig. 4 (bottom). Since there are $2n$ rays, the bound on the number of in-out vertices follows. □

Let Z be an outer region and let n_Z be the number of vertices on Z. If Z is bounded, the complexity of Z is $1 + n_Z$. If Z is unbounded, the complexity of Z is $2 + n_Z$. Hence, the total complexity of the outer zone $\mathcal{Z} = \langle Z_1, \dots, Z_a\rangle$ is $a+n_{\text{i-o}}+n_{\text{i-i}}+n_{\text{o-o}}+a_u$, where $n_{\text{i-o}}$, $n_{\text{i-i}}$, and $n_{\text{o-o}}$ denotes the number of in-out, in-in, and out-out vertices, and a_u is the number of unbounded outer regions. By Claim 3, $n_{\text{i-i}} + n_{\text{o-o}} + a_u \le a$, and by Claim 4, $n_{\text{i-o}} \le 4n$. Thus, the complexity of $\mathcal{Z}$ is at most $2a + 4n \le 8n$, since conv(P) intersects every line of $\mathcal{R}$ as most twice, so $a \le 2n$. □

Theorem 5 *Let $\mathcal{R} = \langle \ell_1, \dots, \ell_n\rangle$ be sequence of n planar lines in general position. Then $c(\mathcal{R}) = O(n^2 137^n)$.*

Proof. Let P be a sequence of points restricted to $\mathcal{R}$ and let $\mathcal{Z} = \{Z_1, \dots, Z_a\}$ be the outer zone of conv(P). To reconstruct the outer zone, it suffices to know (i) the edge e of Z_1 that follows Z_1's supporting arc in clockwise order; and (ii) for $j = 1, \dots, a$, the complexity z_j of Z_j. Indeed, using this information, we can reconstruct the outer zone and obtain a set $\mathcal{C}$ of candidate locations for the convex hull vertices as follows: $\mathcal{C}$ is initialized as the empty set. Starting from e, we walk for $z_1 - 1$ steps in clockwise direction along the boundary of the corresponding cell in $A(\mathcal{R})$ (when taking a step on an unbounded edge of the cell, we proceed to the other unbounded of of the cell). Then we add the current edge e' to $\mathcal{C}$, if $e' \notin \mathcal{C}$, and cross to the neighboring cell of $A(\mathcal{R})$. Next, we continue for $z_2 - 1$ steps in clockwise direction along the boundary of the current cell. After that, we add the current edge into $\mathcal{C}$ if the edge is not contained in $\mathcal{C}$, and change cells in $A(\mathcal{R})$. We continue until we reach the vertex on e again.

To reconstruct conv(P) from the candidate set, we need the information about the vertices of conv(P). Let $\mathcal{C}_\ell = \{e \in \mathcal{C} \mid e \subset \ell\}$, for $\ell = \mathcal{R}$. Since each line $\ell \in \mathcal{R}$ intersects the boundary of conv(P) at most twice, $|\mathcal{C}_\ell| \le 2$. Let e_1 and e_2 be elements of $\mathcal{C}_i$. For $\ell \in \mathcal{R}$, an indicator $b_\ell \in \{1, 2, 3\}$ represents whether (1) conv(P) has a vertex in e_1; (2) conv(P) has a vertex in e_2; or (3) conv(P) has no vertex in ℓ.

Thus, for fixed a, the total number of combinatorially different convex hulls can be estimated as

$$2e_{\mathcal{R}} \cdot 3^n \cdot C_{\mathcal{Z}},$$

where $e_{\mathcal{R}}$ denotes the number of edges of $A(\mathcal{R})$, the second term counts the number of indicator vectors (b_ℓ), and $C_{\mathcal{Z}}$ denotes the number of *complexity vectors* $\langle c_1, \dots, c_a\rangle$. We have $e_{\mathcal{R}} = O(n^2)$. Furthermore, by Lemma 2, the number of complexity vectors is bounded by the number of vectors $(z_1, \dots, z_a)$ with $z_i \in \{2, \dots, n\}$ and $\sum_{i=1}^{a} z_i \le 8n$, which is at most $\binom{8n-a-1}{a-1}$. Thus, for fixed a, the number of combinatorially different convex hulls is at most

$$O\left(n^2 3^n \binom{8n-a}{a}\right).$$

Since this expression grows exponentially with a for $a \in \{1, \dots, 2n\}$, the total number of combinatorially different convex hulls is asymptotically dominated by the term for $a = 2n$, and it is at most

$$O\left(n^2 3^n \binom{6n}{2n}\right) = O\left(n^2 3^n 2^{\frac{6n}{3}\log 3 + \frac{2\cdot 6n}{3}\log\frac{3}{2}}\right)$$
$$= O\left(n^2 3^n 3^{2n} (3/2)^{4n}\right) = O\left(n^2 137^n\right).$$

□

Acknowledgments. This work was begun at the 17th Korean Workshop on Computational Geometry. We would like to thank the organizers and participants for creating a pleasant and productive research environment. W.Mulzer partially supported by DFG grant MU 3501/1. H.Kim supported by the Deutsche Forschungsgemeinschaft within the research training group 'Methods for Discrete Structures' (GRK 1408).

References

[1] K. Buchin, M. Löffler, P. Morin, and W. Mulzer. Preprocessing imprecise points for Delaunay triangulation: Simplified and extended. *Algorithmica*, 61(3):674–693, 2011.

[2] O. Devillers. Delaunay triangulation of imprecise points, preprocess and actually get a fast query time. *J. Comput. Geom. (JoCG)*, 2(1):30–45, 2011.

[3] E. Ezra and W. Mulzer. Convex hull of points lying on lines in $o(n \log n)$ time after preprocessing. *Comput. Geom. Theory Appl.*, 46(4):417–434, 2013.

[4] M. van Kreveld, M. Löffler, and J. S. Mitchell. Preprocessing imprecise points and splitting triangulations. *SIAM Journal on Computing*, 39(7):2990–3000, 2010.

[5] M. Löffler. *Data imprecision in computational geometry.* PhD thesis, 2009.

[6] M. Löffler and J. Snoeyink. Delaunay triangulation of imprecise points in linear time after preprocessing. *Comput. Geom. Theory Appl.*, 43(3):234–242, 2010.