Output-Sensitive Algorithms for Enumerating and Counting Simplices Containing a Given Point in the Plane

Amr Elmasry* Khaled Elbassioni†

Abstract

Given a set of n points $S \subseteq \Re^2$, a specified point $Z \in \Re^2$, it is shown that finding k minimal simplices from S, each of which contains Z, can be done in $O(n+k)$ time. It is also shown that counting the number of all such simplices can be done in $O(n + n\log(k/n+1))$ time, when the number of simplices is k.

1 Introduction

Let $S \subseteq \Re^d$ be a given set of points in the d-dimensional Euclidean space. Given a point $Z \in \Re^d$, let us call a Z-*simplex* any *minimal* subset X of points of S, the convex hull of which contains Z. By Carathéodory theorem (see [8]), the cardinality of any such set X is at most $d+1$. Furthermore, we may assume without loss of generality that Z is the origin. It is a long-standing open problem, known in other polynomially-equivalent forms as the *vertex enumeration* or the *convex hull* problem, to find an algorithm for enumerating all Z-simplices for a given point set $S \subseteq \Re^d$ whose running time is polynomial in $|S|$, d and the number of Z-simplices. See [2, 3, 9].

Since the convex hull problem has been extensively studied, assuming fixed dimension (see [7]), it is natural to ask about the complexity of the enumeration of Z-simplices under the same assumption. In this paper, we consider the case $d = 2$ and present a linear-time algorithm for enumerating such Z-simplices:

Theorem 1 *Given a set of n points $S \subseteq \Re^2$, we can find k Z-simplices from S in time $O(n+k)$.*

We also consider the problem of counting Z-simplices. In [6], it was shown that, for a given set S of n points in the plane, counting the number of triangles with corners at S and containing a point Z can be done in $O(n\log n)$ time. It was furthermore shown in [5] that simultaneously providing such counts for all points in S (each point is considered once as Z) can be done in $O(n^2)$. On the other hand, a lower bound of $\Omega(n\log n)$ for counting the number of such triangles is also known (see [1]). It is straightforward to verify that these bounds can also be extended to the case of counting Z-simplices (with the difference that such simplices can also be just line segments, rather than triangles). We present an output-sensitive algorithm for counting Z-simplices:

Theorem 2 *Given a set of n points $S \subseteq \Re^2$, the Z-simplices from S can be counted in time $O(n + n\log(k/n+1))$, where k is the count of such simplices.*

It remains interesting to investigate whether there is a matching lower bound for the counting problem.

In the next section, we present a simple mapping from our enumeration and counting problems to problems of counting and enumerating certain subsequences of zeros and ones. In Section 3, we show how to enumerate such Z-simplices in linear time as stated by Theorem 1. Finally, an output-sensitive algorithm for counting Z-simplices is given in Section 4.

2 From Z-simplices to sequences of zeros and ones

First, by projecting the points of S on the circumference of the unit circle $C(0,1)$, we can assume without loss of generality that the given set S is located on $C(0,1)$.

Second, let ℓ be an arbitrary diameter of $C(0,1)$, say $\ell = ((-1,0),(1,0))$. Then, the points of S can be represented by a sequence $\mathcal{S} \subseteq \{0,1\}^n$ of zeros and ones of length n in the following way. We begin by replacing each point p below ℓ by $-p$ and mapping it to $\mathcal{M}(p) = 0$. The points p above ℓ remain unchanged and are mapped to $\mathcal{M}(p) = 1$. Now, we imagine that the resulting set of points are sorted by the angle they make with the x-axis. (Note however that the sequence $\mathcal{S}$ is not given in sorted order, but associated with each element $x \in \mathcal{S}$ is its angle $a(x)$ that indicates its relative order in $\mathcal{S}$.) A similar mapping was used by [10].

Now, consider the maximal blocks of consecutive zeros and consecutive ones within the sorted sequence of $\mathcal{S}$, when sorted by angle. We may assume without loss of generality that if two points p_1 and p_2 satisfy $\mathcal{M}(p_1) = 0$, $\mathcal{M}(p_2) = 1$ and have the same angle $a(p_1) = a(p_2)$, then p_1 appears in the sorted sequence before p_2, and consequently they must belong to two consecutive blocks. Let us denote by $B_1^0, B_1^1, B_2^0, B_2^1, \ldots, B_r^0, B_r^1$ the maximal blocks from left

*Department of Computer Engineering and Systems, Alexandria University, Egypt; elmasry@alexeng.edu.eg

†Max-Planck-Institut für Informatik, Saarbrücken, Germany; elbassio@mpi-sb.mpg.de

to right, where B_j^0 denotes a block of zeros, and B_j^1 denotes a block of ones. Let $b_j^0 = |B_j^0|$ and $b_j^1 = |B_j^1|$, for $j \in \{1, \ldots, r\}$, denote the sizes of these blocks, where $b_r^1 \geq 0$, and $b_j^i > 0$ for other values of i and j.

It is now easy to see that Z-simplices from S are in one-to-one correspondence with:

(a) subsequences of $\mathcal{S}$ of length 2 of the form 01, and the angles of the two points coincide (their blocks must be consecutive in the sorted sequence),

(b) subsequences of the sorted sequence of $\mathcal{S}$ of length 3 that have one of the two forms 010, 101 and not containing the subsequences in (a).

The subsequences in (a) are the 1-dimensional Z-simplices representing two points forming a diameter, while those in (b) are the 2-dimensional Z-simplices representing triangles of points containing the center in their interiors. Thus, our problem essentially reduces to finding such subsequences in a given sequence $\mathcal{S} \subseteq \{0, 1\}^n$. We remark that the bound in Theorem 2 also applies to the problem of counting either the diameters or the triangles containing the center, *separately.* On the other hand, the bound in Theorem 1 only applies to the problem of enumerating triangles containing the center, but not to that of enumerating diameters.

3 Enumerating Z-simplices

Given the sequence $\mathcal{S} \subseteq \{0, 1\}^n$ corresponding to the points of S, the first attempt for enumerating all Z-simplices is to sort $\mathcal{S}$ by angle to get the maximal blocks of consecutive zeros and consecutive ones. From such a sorted sequence, it is straightforward to identify any number k among the simplices represented by the subsequences (a) and (b). (Note that k here need not be the total number of Z-simplices.) The time required by this procedure is $O(n \log n + k)$.

Two observations make the bound in Theorem 1 possible. The first is that we do not need to sort the points (zeros and ones) within the blocks; instead we only need to identify which points are in which blocks. The second observation is that when $k = o(n \log n)$ the number of such blocks r is small, and identifying the points in each block can be done in $O(n + k)$. The following two lemmas are used to prove Theorem 1.

Lemma 3 *Given a sequence $\mathcal{S} \subseteq \{0, 1\}^n$ corresponding to points in S, such that the number of the maximal blocks of consecutive zeros in the sorted sequence of $\mathcal{S}$ is r, the number of subsequences of $\mathcal{S}$ representing Z-simplices is $\Omega(nr^2)$ when $r \geq 3$.*

Proof. Let us show that the number of Z-simplices is minimized when each block, but only one, contains exactly one point, and such that all the points mapped to each of two consecutive blocks B_j^0, B_j^1 have the same angle. We start with any sequence $\mathcal{S}$ that has r maximal blocks of consecutive zeros, and transform it through a sequence of steps, that do not increase the number of Z-simplices, to another sequence with the same number of blocks and with the claimed properties. Each such step involves moving the set of points within one or two blocks. We shall make use of the following preposition.

Proposition 4 *Let $N(x, y) = c + c_x x + c_y y + c_{xy} xy$ be a real-valued function in two variables x, y, where c, c_x, c_y, c_{xy} are real constants. Then,*

(i) *if $x \geq 1$, $y \geq 1$ and $c_{xy} \geq 0$, then either $N(x + y - 1, 1) \leq N(x, y)$ or $N(1, x + y - 1) \leq N(x, y)$,*

(ii) *if $x \geq 0$, $y \geq 0$, $c_x \geq c_y$ and $c_{xy} \geq 0$, then $N(0, x + y) \leq N(x, y)$.*

Proof. Let $\delta_1 \geq 0$ and $\delta_2 \geq 0$ and note that

$$\begin{aligned} N(x + \delta_1, y - \delta_1) - N(x, y) &= \\ &- \delta_1 [c_{xy}(\delta_1 + x - y) + c_y - c_x], \quad (1) \\ N(x - \delta_2, y + \delta_2) - N(x, y) &= \\ &- \delta_2 [c_{xy}(\delta_2 + y - x) + c_x - c_y]. \quad (2) \end{aligned}$$

Assume $c_{xy} \geq 0$. Then, it is not possible that both $c_{xy}(\delta_1 + x - y) + c_y - c_x < 0$ and $c_{xy}(\delta_2 + y - x) + c_x - c_y < 0$, for otherwise we get the contradiction $c_{xy}(\delta_1 + \delta_2) < 0$. Thus if $c_{xy}(\delta_1 + x - y) + c_y - c_x \geq 0$, we get by (1) that $N(x + \delta_1, y - \delta_1) \leq N(x, y)$. Otherwise, we get by (2) that $N(x - \delta_2, y + \delta_2) \leq N(x, y)$. Now, (i) follows from the last statement by taking $\delta_1 = y - 1$ and $\delta_2 = x - 1$. Also, (ii) follows from (2) by taking $\delta_2 = x$. □

Let μ_j^0 and μ_j^1, for $j \in \{1, \ldots, r\}$, be the count of zeros and ones with the same angle from two consecutive blocks of the sorted sequence of $\mathcal{S}$. (Note that by our assumption on the sorted sequence, $\mathcal{S}$ does not contain a 1 followed by 0, for which the corresponding points have the same angle.) Let us further denote by η_j^i the difference $b_j^i - \mu_j^i$, for $j \in \{1, \ldots, r\}$ and $i \in \{0, 1\}$. For a subsequence $\mathcal{S}'$ of zeros and ones, $j, j' \in \{1, \ldots, r\}$, and $i, i' \in \{0, 1\}$, denote by $N_{\mathcal{S}'}(B_j^i, B_{j'}^{i'})$ the number of occurrences of $\mathcal{S}'$ in the subsequence of $\mathcal{S}$ starting from block B_j^i and ending by block $B_{j'}^{i'}$.

Now, fix the number of points in each of the blocks. Consider a block B_j^i, and assume without loss of generality that $i = 0$. The case for $i = 1$ is symmetric. Then, the number of Z-simplices, including both the 1-dimensional and 2-dimensional ones, can be written as a function of η_j^0 and μ_j^0 as follows:

$$
\begin{aligned}
N(\eta_j^0,\mu_j^0) &= c + N_{01}(B_1^0, B_{j-1}^1)(\eta_j^0+\mu_j^0) \\
&\quad +(\mu_j^0+\eta_j^0)N_{10}(B_{j+1}^1, B_r^0) \\
&+\eta_j^0(\mu_j^1+\eta_j^1)N_0(B_{j+1}^0, B_r^0) + \mu_j^0\eta_j^1 N_0(B_{j+1}^0, B_r^0) \\
&\quad +N_1(B_1^1, B_{j-1}^1)\eta_j^0(\eta_j^1+\mu_j^1+N_1(B_{j+1}^1, B_r^1)) \\
&\quad +N_1(B_1^1, B_{j-1}^1)\mu_j^0(\eta_j^1+N_1(B_{j+1}^1, B_r^1)) + \mu_j^0\mu_j^1, \qquad (3)
\end{aligned}
$$

where the value of c does not depend on η_j^0 or μ_j^0. We now apply Proposition 4 with $x \leftarrow \eta_j^0$ and $y \leftarrow \mu_j^0$, where $c_{xy} = 0$ and

$$
c_x - c_y = \mu_j^1[N_0(B_{j+1}^0, B_r^0) + N_1(B_1^1, B_{j-1}^1) - 1].
$$

Since $r \geq 3$, we have $N_0(B_{j+1}^0, B_r^0) + N_1(B_1^1, B_{j-1}^1) > 1$, and thus Proposition 4 (ii) implies that setting $\eta_j^0 \leftarrow 0$ and $\mu_j^0 \leftarrow \mu_j^0 + \eta_j^0$ will not increase $N(\eta_j^0, \mu_j^0)$. In other words, by changing the angles of the points in block B_j^0 so that they all make the same angle as that of the first point in Block B_j^1 (if such a block exists), we do not increase the total number of Z-simplices. Hence, we may assume from now on that $\eta_j^0 = 0$ for all $j = 1,\ldots,r$. Similarly, by setting $i = 1$, we get $\eta_j^1 = 0$ for all $j = 1,\ldots,r$. Thus, each two consecutive blocks B_j^0, B_j^1 are associated with a single angle to which all the points of the two blocks are mapped.

Let B_j^i and $B_{j'}^{i'}$ be two distinct blocks. As in (3), we write the number of Z-simplices in terms of μ_j^i and $\mu_{j'}^{i'}$

$$
N(\mu_j^i, \mu_{j'}^{i'}) = c_0 + c_1\mu_j^i + c_2\mu_{j'}^{i'} + c_3\mu_j^i\mu_{j'}^{i'},
$$

where c_0, c_1, c_2 and c_3 have *non-negative* values that depend on neither μ_j^i nor $\mu_{j'}^{i'}$. Now, let us apply Proposition 4 (i) with $x \leftarrow \mu_j^i$ and $y \leftarrow \mu_{j'}^{i'}$. It follows that we can assume that either $\mu_j^i = 1$ or $\mu_{j'}^{i'} = 1$. In other words, there is a choice of indices, say i and j, such that if we move all but one point of (the set of points currently mapped to) B_j^i until they are mapped to $B_{j'}^{i'}$ with the same angle associated with the points currently in $B_{j'}^{i'}$, we do not increase the total number of Z-simplices. This establishes our claim.

What remains is to count the number of Z-simplices in such a case, which is a lower bound on the possible number of Z-simplices. Let B_j^i be the block that has $n-2r+1$ points, and call this block the *long block.* We consider the case when $i = 1$ and $b_r^1 > 0$; the other cases are similar. First, consider the count of Z-simplices that have one point from the points of this long block B_j^1. There are $\Theta(r^2)$ possibilities for selecting two other points from another two blocks. This follows from the fact that after selecting two different indices out of the possible $r-1$ indices, other than j, the type (zeros or ones) of the blocks to be selected are imposed. (For example, if the two selected indices are smaller than j, the type of the block with the smaller index must be a 1 and the other block must be a 0. The other cases are similar.) Since $r \geq 3$, this accounts for a total count of $(n-2r+1)\Theta(r^2)$ for such Z-simplices. Next, consider the Z-simplices that have none of the three points from the long block. There are $\Theta(r^3)$ possibilities for selecting these points from three blocks other than B_j^i. This is done by selecting three different indices out of the possible $r-1$ indices, other than j. For each such triple of indices, there is two possible choices, either 010 or 101. Also, there are the $n-r$ possible 1-dimensional Z-simplices. The total count of Z-simplices in this case is, therefore, $(n-2r+1)\Theta(r^2) + \Theta(r^3) + (n-r)$. If $r = \Theta(n)$, the bound $\Theta(r^3)$ and the fact that $n \geq 2r$ imply that the number of such Z-simplices is $\Theta(nr^2)$. If $r = o(n)$, the term $(n-2r+1)\Theta(r^2)$ implies the same bound of $\Theta(nr^2)$. It follows that the number of Z-simplices for any configuration of points is $\Omega(nr^2)$. □

Lemma 5 *Given a sequence $\mathcal{S} \subseteq \{0,1\}^n$ corresponding to points in S, such that the number of the maximal blocks of consecutive zeros in the sorted sequence of $\mathcal{S}$ is r, the elements of each of these blocks can be identified in time $O(nr)$.*

Proof. To identify which points in $\mathcal{S}$ are in which blocks we proceed as follows. First, $\mathcal{S}$ is divided into a subsequence of zeros and another subsequence of ones. The point that has the smallest angle among the subsequence of ones is identified, and the subsequence of zeros is partitioned into two subsequences with respect to the angle of this identified one. The subsequence of zeros that have a smaller angle is the first block of zeros. This step is done in $O(n)$. We proceed with the remaining points among the subsequence of zeros by identifying the point with the smallest angle. The subsequence of ones is partitioned into two subsequences with respect to the angle of this identified zero. The subsequence of ones that have a smaller angle is the first block of ones. This partitioning step is repeated alternatively between zeros and ones, until all the blocks are identified. Since the number of such blocks is r and each step of identifying a block requires $O(n)$, the lemma follows. □

We are now ready to complete the proof of Theorem 1. Given the sequence $\mathcal{S} \subseteq \{0,1\}^n$ corresponding to the points of S, we apply the algorithm in Lemma 5 to identify the r blocks of consecutive zeros and consecutive ones in $O(nr)$ time. Subsequently, finding all the Z-simplices is straightforward and can be done in $O(n)$ time. If $r \leq 2$ we spend $O(n)$ time to find the Z-simplices or to realize that there are none. If $r \geq 3$, using Lemma 3, the time spent by the algorithm is $O(k)$, where k is the number of the identified Z-simplices.

We can also apply the algorithm in Lemma 5 to find a specified number of the Z-simplices. Given an integer

k representing the required number of Z-simplices to be found, decide a value for $r \geq 3$ such that $r = \frac{ck}{n}$ for some constant c . We apply the partitioning step of the algorithm in Lemma 5 at most r times to identify the first at most r blocks of consecutive zeros and consecutive ones within $\mathcal{S}$. By using an adequate value for the constant c, Lemma 3 implies that we get the required k Z-simplices, or otherwise all of them, in $\Theta(n+k)$.

4 Counting Z-simplices

As pointed out in Section 3, we do not need to sort $\mathcal{S}$. Instead, we only need to identify the consecutive blocks of zeros and ones within the sorted sequence of $\mathcal{S}$ (i.e. which points are in which blocks). Then, we need to count the number of subsequences of the form 010 and 101 in such sequence. Such a count can be easily obtained in $O(n)$ time once the blocks of zeros and ones are identified, as illustrated by the next lemma. See [5, 6] for more details.

Lemma 6 *Given a sequence of alternating consecutive blocks of zeros and ones, the number of subsequences of the form* 010 *and* 101 *in the given sequence can be counted in $O(n)$ time.*

Proof. Let us, without loss of generality, count the number of 010 subsequences. To do this, we scan the given sequence from left to right, computing the sizes of maximal blocks of consecutive zeros and consecutive ones b_j^i, for all $j \in \{1, \ldots, r\}$ and $i \in \{0, 1\}$. Now, the number N_{010} of subsequences of 010 is given by the following formula

$$N_{010} = \sum_{1 \leq j_1 < j_2 < j_3 \leq r} b_{j_1}^0 b_{j_2}^1 b_{j_3}^0. \tag{4}$$

Such a formula (4) can be computed in linear time. For each $j = 2, \ldots, r-1$, incrementally compute the prefix and postfix sums:

$$\alpha_1^0 = b_1^0, \qquad \alpha_j^0 = \alpha_{j-1}^0 + b_j^0,$$
$$\beta_1^0 = \sum_{h=2}^{r} b_h^0, \qquad \beta_j^0 = \beta_{j-1}^0 - b_j^0.$$

It follows that $N_{010} = \sum_{j=1}^{r-1} \alpha_j^0 b_j^1 \beta_j^0$. Similarly, the numbers α_j^1, β_j^1 can be defined, for $j = 1, \ldots, r-1$, from which N_{101} can be computed.

To account for the 1-dimensional Z-simplices, we need to count the subsequences of consecutive zeros and ones with the same angle. As defined above, let μ_j^0 and μ_j^1, for $j = 1, \ldots, r$, be the count of zeros and ones with the same angle from two consecutive blocks of the sorted sequence of $\mathcal{S}$. (Note that by our assumption on the sorted sequence, $\mathcal{S}$ does not contain a 1 followed by 0 for which the corresponding points have the same angle.) The count of 1-dimensional Z-simplices is, therefore, $\sum_{j=1}^{r} \mu_j^0 \mu_j^1$. Note that we still need to adjust the count N_{010} of the 2-dimensional Z-simplices to account for those simplices that contain 1-dimensional Z-simplices as subsets. It is not difficult to see that the adjusted count of N_{010} is $\sum_{j=1}^{r-1} (\alpha_j^0 b_j^1 \beta_j^0 - \mu_j^0 \mu_j^1 \beta_j^0)$. The adjusted count of N_{101} can be calculated similarly. □

Applying the algorithm of Lemma 5, we have an output-sensitive algorithm that counts Z-simplices in $O(nr)$. Since, using Lemma 3, $r = O(\sqrt{k/n})$, it follows that such an algorithm can be used to count the number of Z-simplices in $O(n + \sqrt{kn})$, where k is the output of the algorithm. This motivates improving the algorithm, and hence the bound, of Lemma 5. (The $O(nr)$ bound of Lemma 5 is enough for the enumeration algorithm.)

We use any of the algorithms in [4] to achieve the next lemma, which implies an algorithm for counting Z-simplices in $O(n + n \log(k/n + 1))$.

Lemma 7 *Given a sequence $\mathcal{S} \subseteq \{0,1\}^n$ corresponding to points in S, such that the number of the maximal blocks of consecutive zeros in the sorted sequence of $\mathcal{S}$ is r, the elements of each of these blocks can be identified in time $O(n \log r)$.*

References

[1] G. Aloupis, C. Corts, F. Gmez, M. Soss and G. Toussaint, Lower bounds on computing statistical depth, *Computational Statistics and Data Analysis*, **40**(2):(2002), 223-229.

[2] E. Boros, K. Elbassioni, V. Gurvich and L. Khachiyan, On enumerating minimal dicuts and strongly connected subgraphs, *10th Conf. on Integer Programm. and Combinatorial Optimization*, LNCS 3064 (2004), 152-162.

[3] M. Bussieck and M. Lübbecke, The vertex set of a 0/1-Polytope is Strongly P-Enumerable, *Computational Geometry: Theory and App.* **11**(2):(1998), 103-109.

[4] A. Elmasry, Distribution-sensitive set multi-partitioning, *1st International Conference on the Analysis of Algorithms*, (2005).

[5] S. Khuller and J. S.B. Mitchell, On a triangle counting problem, *Info. Proc. Letters*, **33**:(1989/90), 319-321.

[6] P. Rousseeuw and I. Ruts, The depth function of a population distribution, *Metrika*, **49**(3):(1999), 213-244.

[7] R. Seidel, Convex hull computations, *Handbook of Discrete and Computational Geometry, J. E. Goodman and J. O'Rourke, eds.*, CRC Press LLC, (1997), 361-377.

[8] A. Schrijver. *Theory of Linear and Integer Programming*, Wiley-Interscience, (1986).

[9] G. Swart, Finding the convex hull facet by facet, *Journal of Algorithms*, **6**:(1985), 17-48.

[10] M. Szegedy and J. Pach, The number of simplices embracing the origin, in: Discrete Geometry: In honor of W. Kuperberg's 60th Birthday (A. Bezdek, ed.), *Pure and Applied Mathematics*, (2003), 381-386.